Stochastic logistic model with environmental noise

Brian Dennis
Dept Fish and Wildlife Resources
and
Dept Statistical Science
University of Idaho

Questions

I. What plausible biological mechanisms, if any, lead to logistic growth?

Derivations of & mechanisms behind logistic

II. How can the logistic be made more useful?

Accounting for stochastic forces $\Leftrightarrow$ Fitting to data and evaluating

$$\frac{dN}{dt} = rN - \left(\frac{r}{\overline{N}}\right)N^2 = aN - bN^2$$

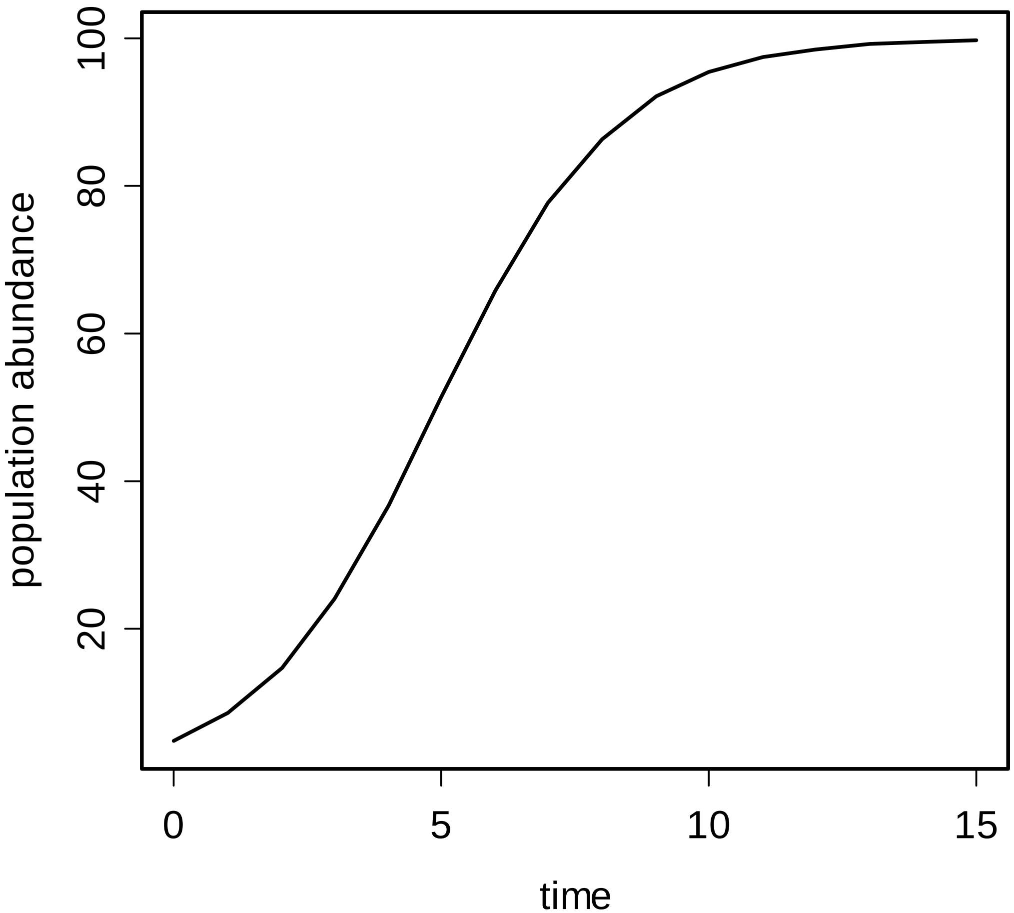
N: population abundance at time t

r: maximum per-unit-abundance growth rate

$\overline{N}$: equilibrium abundance

transformation: $Y = \frac{1}{N}$ yields linear DE for Y

solution trajectory: $N = \frac{\overline{N}}{1 + \left(\frac{\overline{N} - N_0}{N_0}\right)e^{-rt}}$


population abundance
100
80
60
40
20
0
5
10
15
time

I. Derivations/mechanisms

A. Verhulst (1838) "logistique”
   added $-bN^2$ and $-bN^\alpha$ to the exponential growth DE as an adjustment

B. Lotka (1925)

$$\frac{dN}{dt} = m(N) \approx aN + bN^2$$

Handwaves about the sign of b

Taylor series expands around a particular point— what point does he have in mind?

Lotka's derivation is parroted in many ecology textbooks and websites (Hutchinson 1978)

C. Chapman (1928)

$$\frac{dN}{dt} = rN \times (\text{“environmental resistance”}) = rN\left(1 - \frac{N}{K}\right)$$

D. Dennis and Patil (1984), Dennis (1989a)   Re-formulation of Lotka

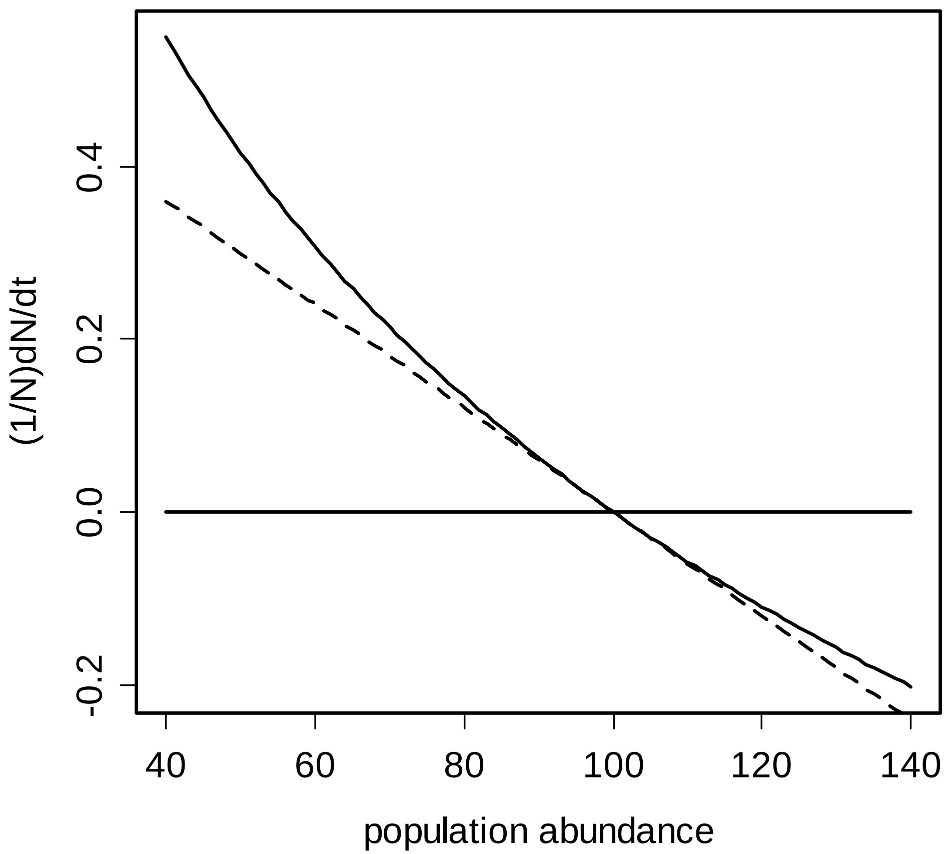
$\frac{dN}{dt} = Ng(N)$ $g(N)$ is per-unit-abundance growth rate,
presumed decreasing

$$g(\overline{N}) = 0, \quad g'(\overline{N}) < 0$$

Near $\overline{N}$,

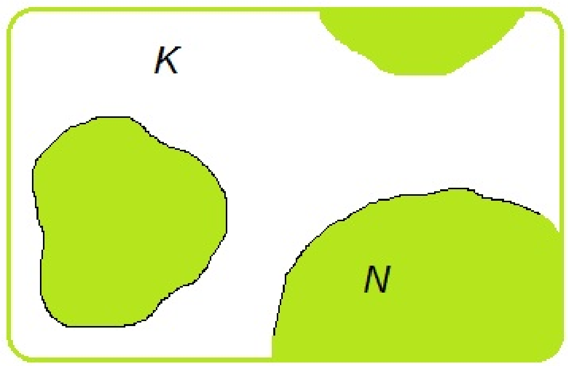
$$g(N) \approx g(\overline{N}) + (N - \overline{N})g'(\overline{N}) = a - bN$$

$$a = -\overline{N}g'(\overline{N}) \qquad b = -g'(\overline{N})$$



(1/N)dN/dt
0.4
0.2
0.0
-0.2
40
60
80
100
120
140
population abundance

E. Substrate-product model: "something replacing something else"



Invasions: cheat grass replacing native grassland. Epidemics: infecteds replacing susceptibles. Commerce: Wal-Mart replacing K-Mart. Innovation: DVDs replacing VHSs. Dissemination of ideas:

Like an autocatalytic chemical reaction (Pearl and Reed 1920, Reed and Berkson 1929):

$$\frac{dN}{dt} = k_1 N(K - N)$$

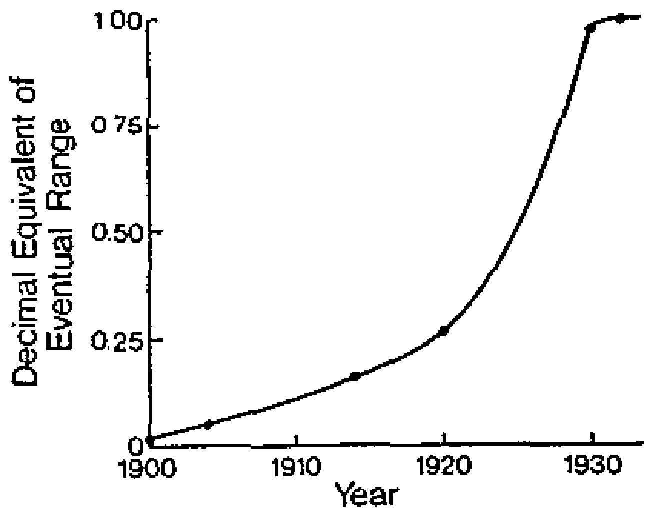


Fig. 2. Progress of *B. tectorum* as a plot of area occupied versus time based on accumulative percentage of eventual range occupied by 1900, 1904, 1914, 1920 and 1930.

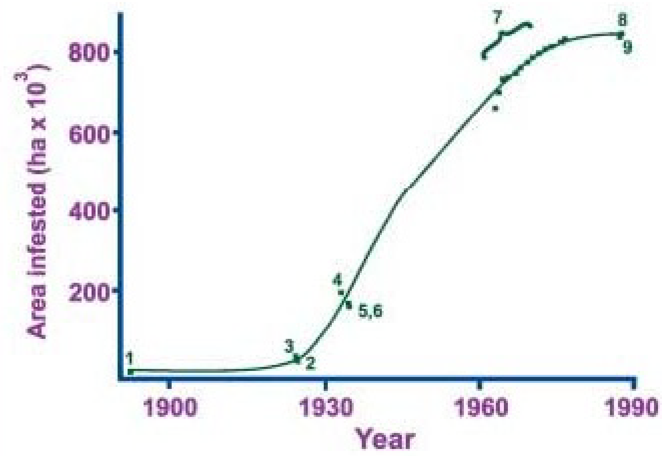


Figure 2 - Many invaders occupy new ranges at an accelerating rate with pronounced "lag" and "log" phases of proliferation and spread. This initial slow rate of range occupation may be indistinguishable from the rate of spread displayed by non-invasive (but nevertheless non-indigenous) species in a new range, thus hampering the early identification of future invaders. Terrestrial plant invasions most commonly illustrate this pattern (e.g. the spread of *Opuntia aurantiaca* in South Africa) (Moran and Zimmerman 1991 and sources [numbers 1-9] therein). By contrast, invaders in other taxonomic groups may show no lag in range expansion and rapidly occupy new range upon entry.

F. (Williams 1972) Explicit limiting nutrient (or substrate)

1. Closed system

$$C \longrightarrow N$$

$$\frac{dC}{dt} = -\alpha k_1 C N$$

α: conversion (in)efficiency

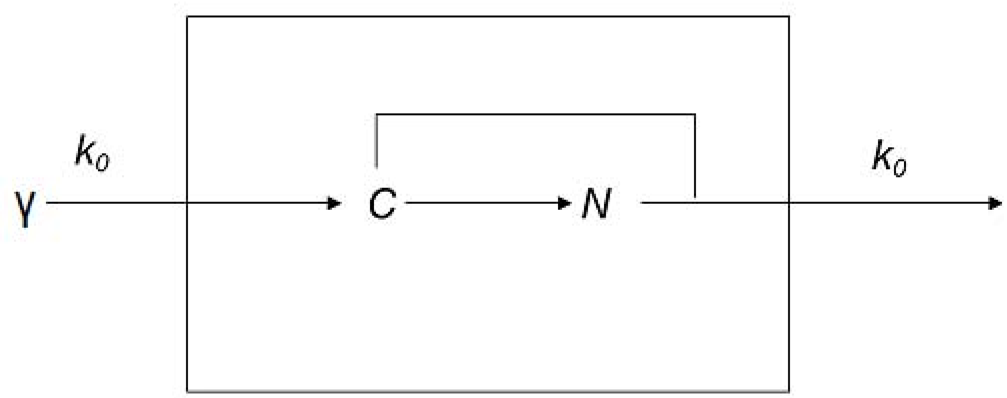
$$\frac{dN}{dt} = k_1 C N$$

mass balance: $C + \alpha N = C_0 + \alpha N_0$

$$\Rightarrow \; C = C_0 + \alpha N_0 - \alpha N$$

$$\frac{dN}{dt} = rN - \frac{r}{\overline{N}} N^2 \qquad r = k_1(C_0 + \alpha N_0), \qquad \overline{N} = \frac{1}{\alpha}(C_0 + \alpha N_0)$$

2. Open system (like a laboratory microbial chemostat)



$\frac{dC}{dt} = k_0\gamma - \alpha k_1 CN - k_0 C$

$\frac{dN}{dt} = k_1 CN - k_0 N$

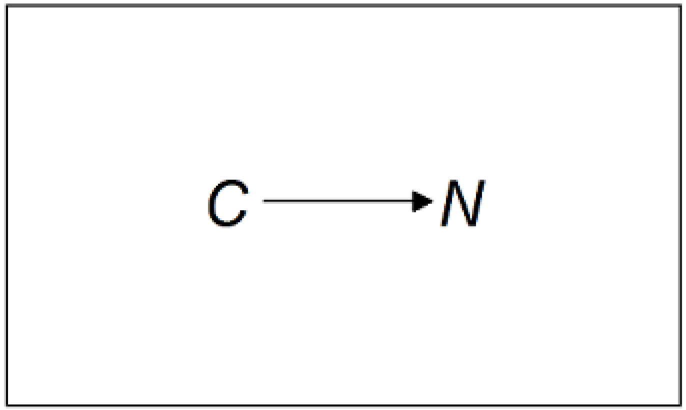
γ: nutrient concentration in inflow
k_0: dilution rate

mass balance: $C + \alpha N = \gamma + (C_0 + \alpha N_0 - \gamma)e^{-k_0 t}$

$\Rightarrow \frac{dN}{dt}$ is a logistic with time-dependent coefficients (Dennis 1978),

is exactly logistic if $\gamma = C_0 + \alpha N_0$, and becomes logistic as $t \to \infty$

3. Closed system, Michaelis-Menton nutrient uptake



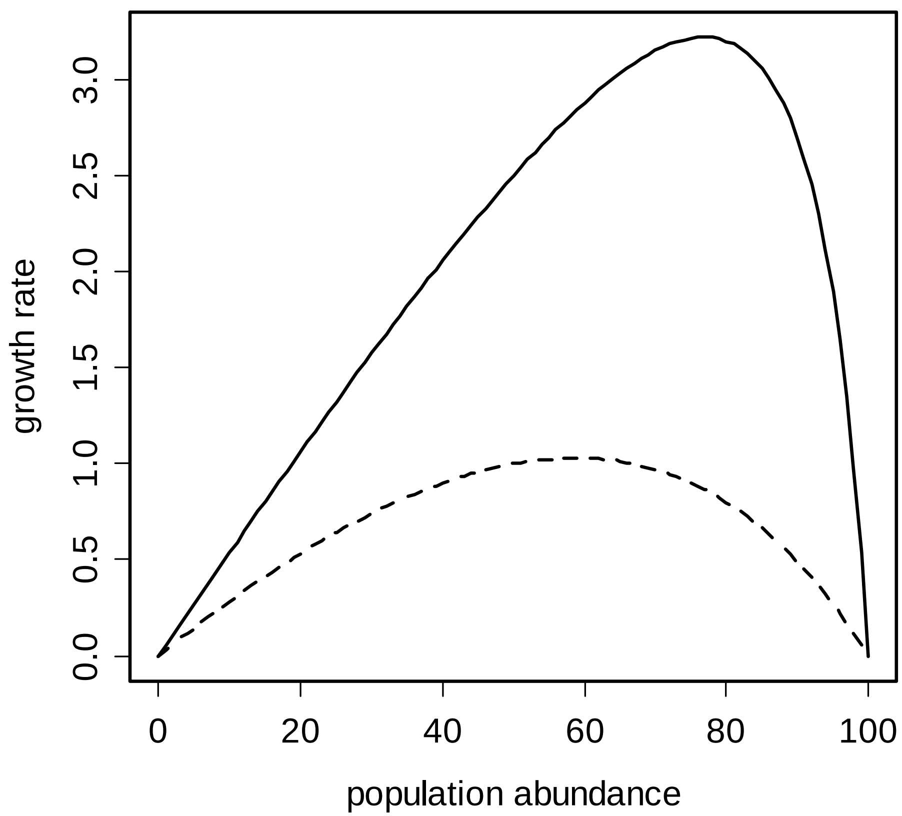
$C + V \rightleftharpoons C{\sim}V \rightarrow V + N$

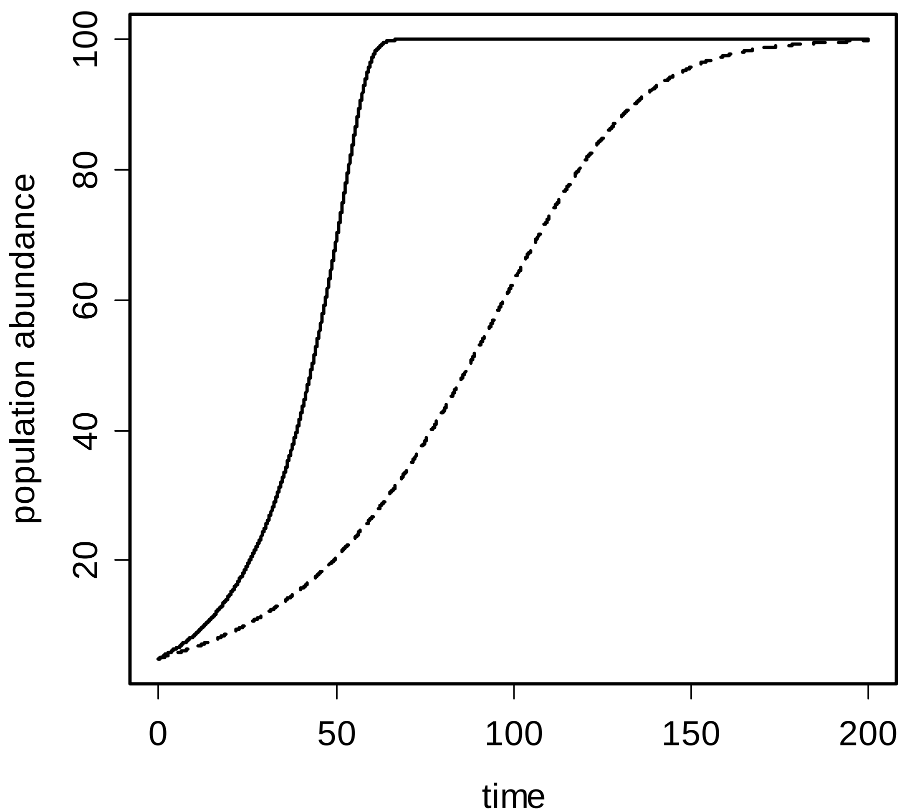
V: uptake bottleneck (gut volume, cell membrane sites, filter surface, etc.)

$\Rightarrow$ rate of product formation per unit abundance $\approx \frac{FC}{K+C}$
(analogous to Michaelis-Menten enzyme kinetics)

$$\frac{dC}{dt} = -\alpha \frac{FCN}{K+C}$$

$$\frac{dN}{dt} = \frac{FCN}{K+C}$$


growth rate
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0
20
40
60
80
100
population abundance



population abundance
100
80
60
40
20
0
50
100
150
200
time

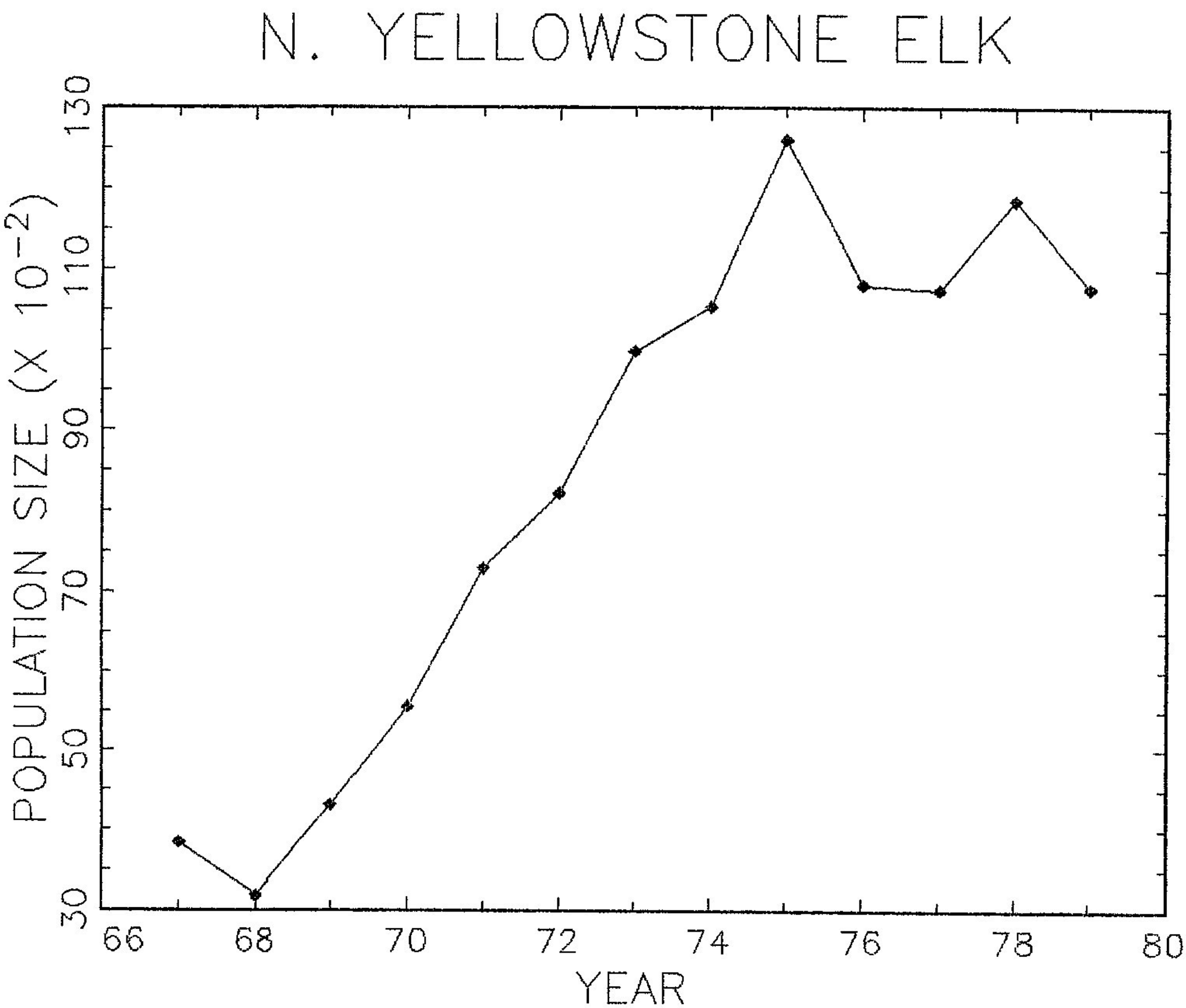
II. Stochastic forces and data analysis

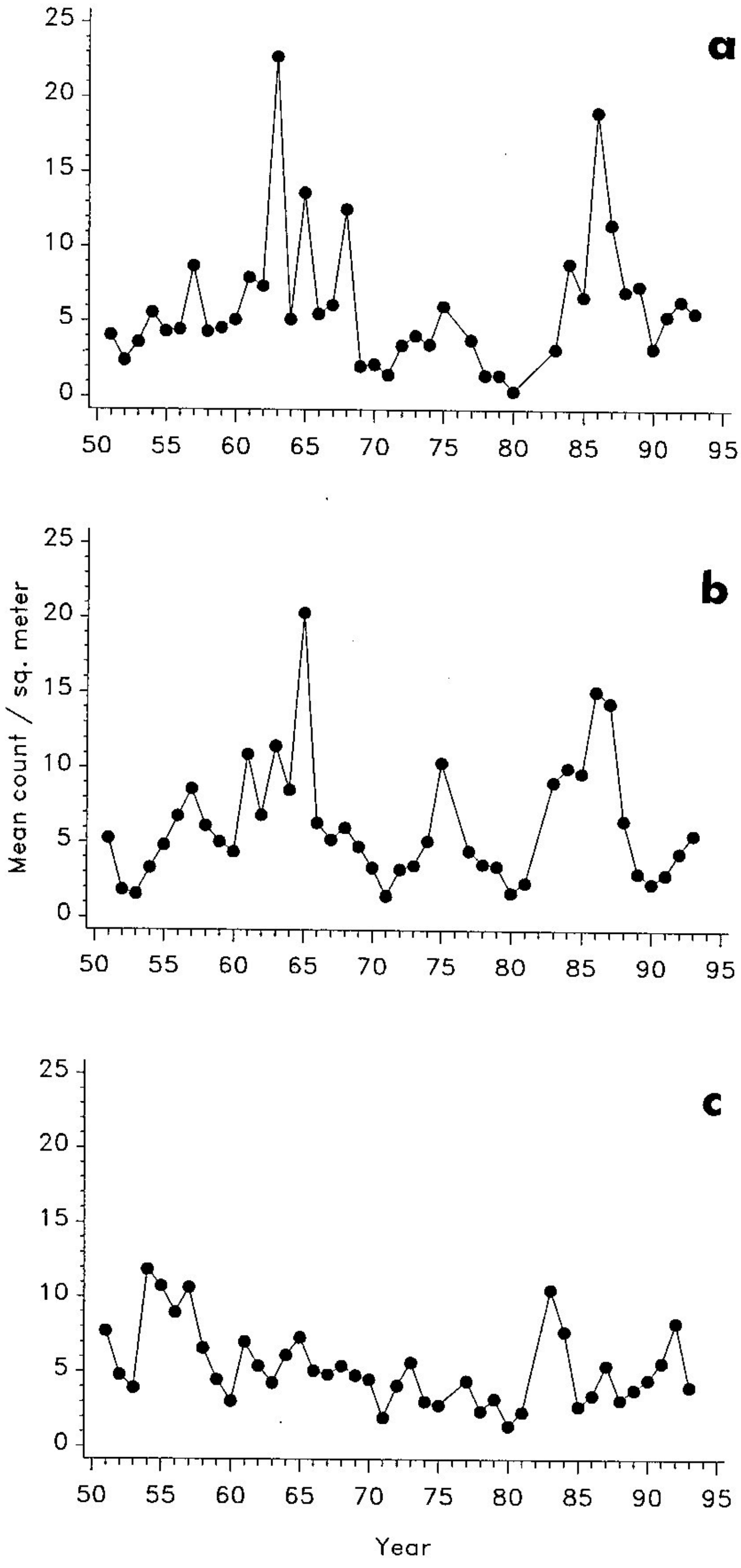
A. Problems with using the logistic as is.

1. Life is stochastic (data depart from model in noisy ways)

2. How to do inferences (estimation, evaluation)

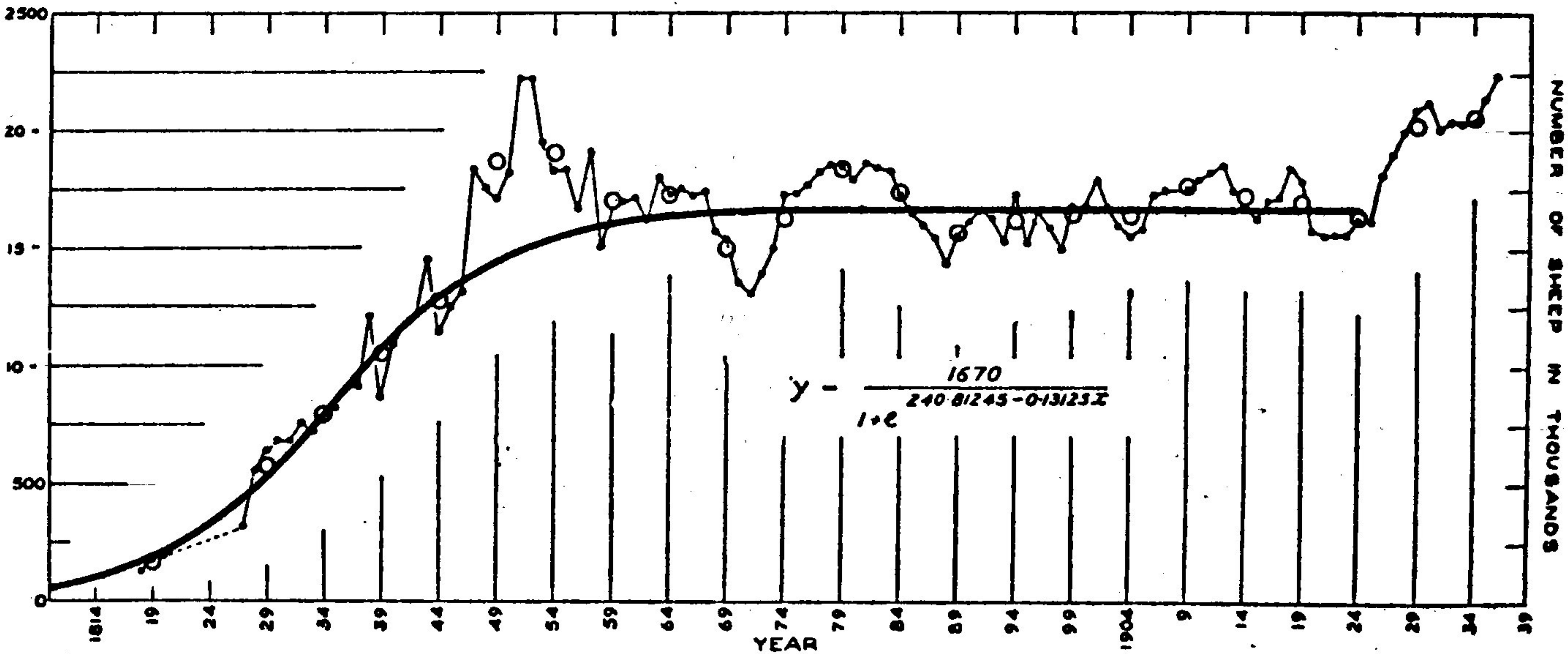
3. Modeling objectives might require more realistic modeling


N. YELLOWSTONE ELK
POPULATION SIZE (X 10^{-2})
130
110
90
70
50
30
66
68
70
72
74
76
78
80
YEAR



a
b
c
Mean count / sq. meter
Year
0
5
10
15
20
25
50
55
60
65
70
75
80
85
90
95

GROWTH OF SHEEP POPULATION IN TASMANIA



B. Off-the-shelf solution: least squares (Leach 1981)

Data: time series population abundances

0	t_1	t_2	$\cdots$	t_q
n_0	n_1	n_2	$\cdots$	n_q

Model parameters: r, $\overline{N}$, σ^2

Minimize departures of observations from logistic trajectory:

$$SS(r, \overline{N}, \sigma^2) = \sum_{i=1}^{q} \left(n_i - \frac{\overline{N}}{1 + \left(\frac{\overline{N} - n_0}{n_0}\right) e^{-rt_i}} \right)^2$$

This procedure implicitly assumes a stochastic logistic model in which the underlying growth is deterministic and the departures from the trajectory are due to observation or measurement error.

Ecological populations themselves are stochastic (process noise).

C. Models of process noise

Seek: a useful, general stochastic version of

$$dN = m(N)dt$$

Answer: a diffusion process is in the form

$$dN_t = m(N_t)dt + \sqrt{v(N_t)}\, dW_t,$$

where:

N_t is population abundance at time t,
$dW_t \sim$ normal(0, dt)
$m(n)$ is the infinitesimal mean ("skeleton")
$v(n)$ is the infinitesimal variance

D. Stochastic logistic with “environmental noise”:

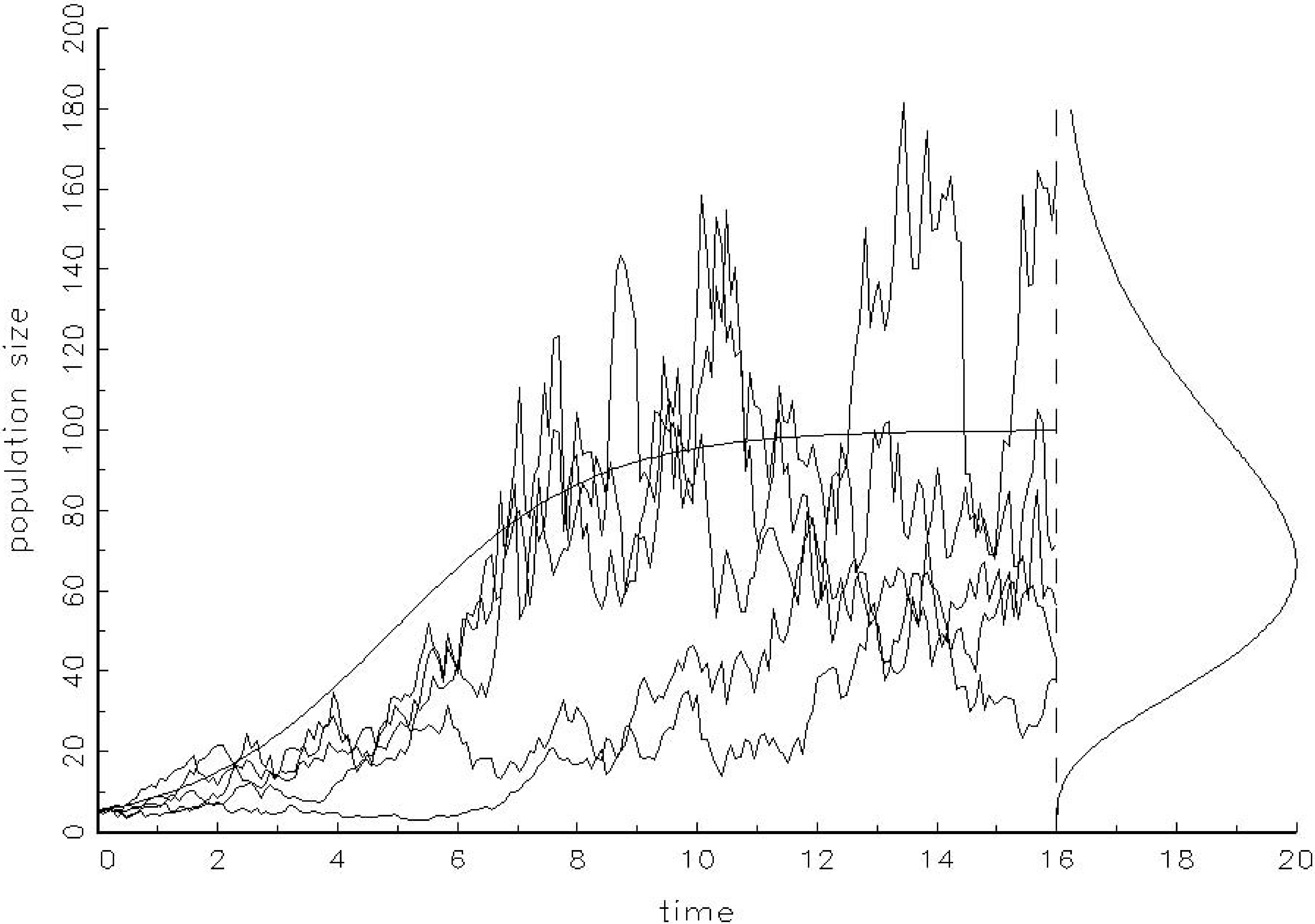
$$m(n) = rn + \left(\frac{r}{k}\right)n^2$$

$$v(n) = \sigma^2 n^2$$

(I'll use "k" instead of $\overline{N}$ from here on... "equilibrium" of the stochastic model will be seen to be a probability distribution)

this form of $v(n) \Rightarrow$ ln(N_t) has a constant infinitesimal variance

$\Rightarrow$ substantial fluctuations at all population sizes, not just at low population sizes



population size
time
0
20
40
60
80
100
120
140
160
180
200
0
2
4
6
8
10
12
14
16
18
20

Transition probability density function: $p(n, t \mid n_0)$

Pr$[a < N_t \leq b] = \int_a^b p(n, t \mid n_0) dn$

$$\int_0^\infty p(n, t \mid n_0) dn = 1$$

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 [vp]}{\partial n^2} - \frac{\partial [mp]}{\partial n}$$

(Fokker-Planck equation)

The FP for the stochastic logistic has been solved!

$$p(n, t \mid n_0) = p(n) \sum_{\lambda \epsilon A} \phi_\lambda(n) \phi_\lambda(n_0) e^{-\lambda t}$$

ϕ_λ is related to a confluent hypergeometric function

Stationary pdf of a diffusion process:

$$p(n, t \mid n_0) \underset{\text{as } t \to \infty}{\longrightarrow} p(n)$$

$$p(n) = \frac{C}{v(n)} e^{2\int \frac{m(n)}{v(n)} dn}$$

stationary pdf exists if

$$\int_0^\infty p(n) dn = 1$$

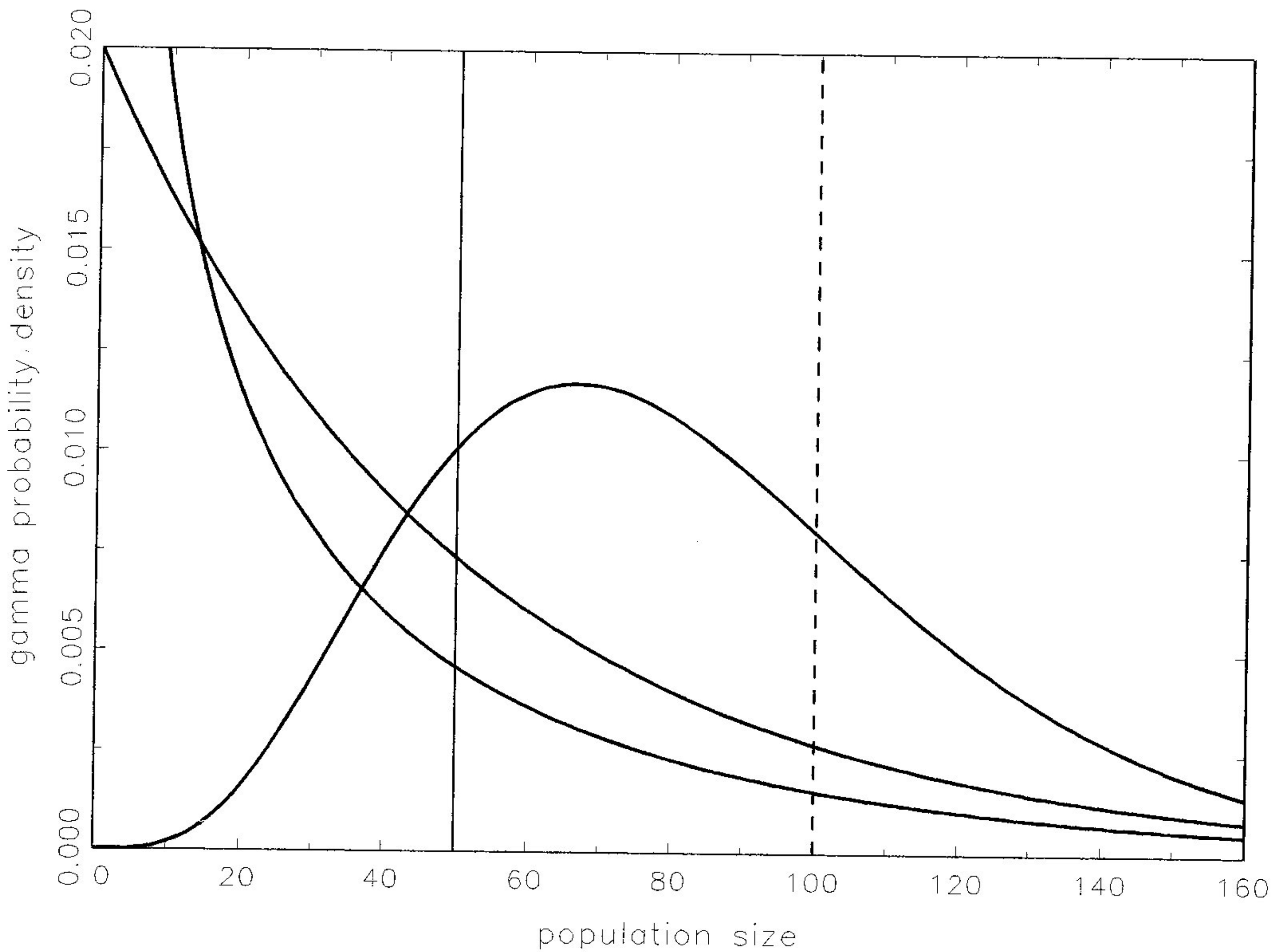
Stationary pdf for stochastic logistic:

$$p(n) = \frac{\beta^\alpha}{\Gamma(\alpha)} n^{\alpha-1} e^{-\beta n}$$

(gamma distribution)

$$\alpha = \frac{2r}{\sigma^2} - 1, \;\; \beta = \frac{2r}{k\sigma^2}$$

stationary pdf exists if $\alpha > 0$


gamma probability density
0.000
0.005
0.010
0.015
0.020
0
20
40
60
80
100
120
140
160
population size

Moments of N_t:

$$u_\nu(n_0, t) \;=\; \mathsf{E}[(N_t)^\nu \mid N_0 = n_0]$$

$$\frac{\partial u_\nu}{\partial t} \;=\; \frac{v(n_0)}{2}\frac{\partial^2 u_\nu}{\partial n_0^2} \;+\; m(n_0)\frac{\partial u_\nu}{\partial n_0}$$

(backward equation)

$$u_\nu(n_0, 0) \;=\; n_0^\nu$$

Approximation for the transition pdf (Dennis 1989b)

1. Find approximations for $u_1(n_0, t)$ and $u_2(n_0, t)$

2. Use a gamma pdf with those moments

1. Singular perturbation (small noise expansion) of backward equation: write

$$u_\nu(n_0,t) = w_0(n_0,t) + \\ \sigma w_1(n_0,t) + \sigma^2 w_2(n_0,t) + \cdots$$

substitute into the PDE, equate coefficients of like powers of σ, solve resulting PDEs for w_0, w_1, ..., subject to initial condition (laborious)

Wieszak (1988) proved that this perturbation is asymptotically correct

$$u_1(n_0,t) \;\approx\; k\Big/\left[1+\left(\frac{k-n_0}{n_0}\right)e^{-rt}\right]$$
$$+\;\frac{k\sigma^2}{2r}\left\{\left[1-\frac{2k}{n_0}\right]e^{-2rt} \;+\; \frac{2k}{n_0}e^{-rt} \;-\; 1 \;-\; 2r\left[\frac{k-n_0}{n_0}\right]te^{-rt}\right\}$$
$$\times\;\left[1+\left(\frac{k-n_0}{n_0}\right)e^{-rt}\right]^{-3}$$

$$u_2(n_0,t) \;\approx\; k^2\Big/\left[1+\left(\frac{k-n_0}{n_0}\right)e^{-rt}\right]^2$$
$$+\;\left(\frac{\sigma^2}{2}\right)\left(\frac{k}{r}\right)^2 2r\left\{\left[-\frac{4k}{n_0}+\frac{5}{2}\right]e^{-2rt} \;+\; \left[\frac{4k}{n_0}-2\right]e^{-rt}-\frac{1}{2}\right.$$
$$\left.+\;r\left(\frac{k-n_0}{n_0}\right)te^{-rt}\left[\left(\frac{k-n_0}{n_0}\right)e^{-rt}-2\right]\right\}\left[1+\left(\frac{k-n_0}{n_0}\right)e^{-rt}\right]^{-4}$$

2. Use a gamma distribution w/ matching moments as transition pdf:

$$p(n,t\,|\,n_0) \;\approx\; \frac{\beta_t}{\Gamma(\alpha_t)}n^{\alpha_t-1}e^{-\beta_t n}$$

$$\alpha_t \;\equiv\; \alpha(n_0,t) \;=\; \frac{u_1^{\,2}}{u_2-u_1^{\,2}}$$

$$\beta_t \;\equiv\; \beta(n_0,t) \;=\; \frac{u_1}{u_2-u_1^{\,2}}$$

E. Simulations: does the approximation work?

$$dN_t = \left[rN_t + \left(\frac{r}{k}\right){N_t}^2\right]dt + \sigma N_t dW_t$$

1000 realizations ending at each time t ($t =$ 1, 2, 3, ...until stationarity)

Density probability plot:

excellent graphical tool for assessing goodness of fit for a continuous distribution (related to the probability plot)

Y, r.v. with continuous distribution

$F(y)$, $f(y)$: cdf & pdf models to be assessed

$y_{(1)}$, $y_{(2)}$, ..., $y_{(q)}$ data (ordered)

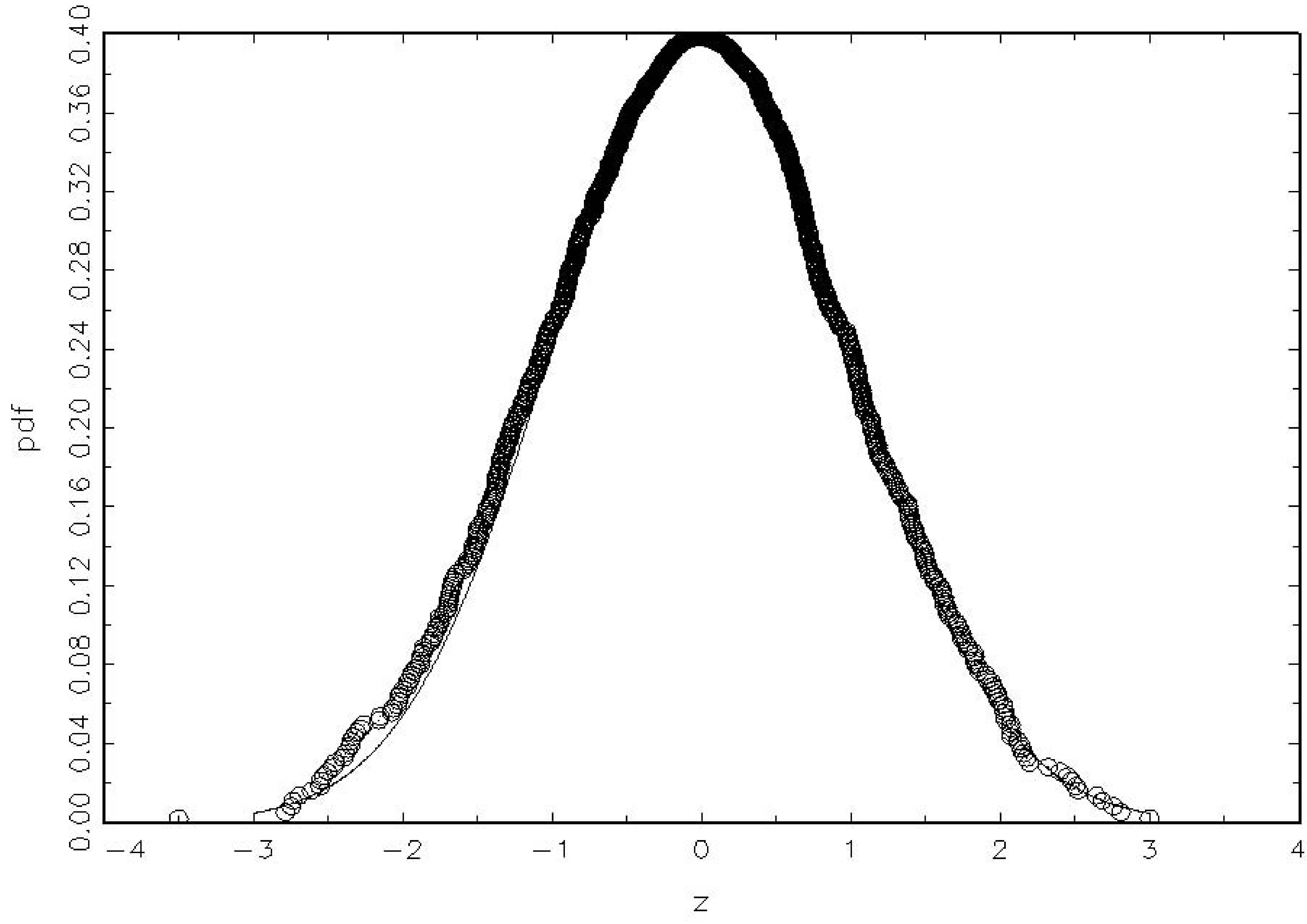
usual "probability plot":

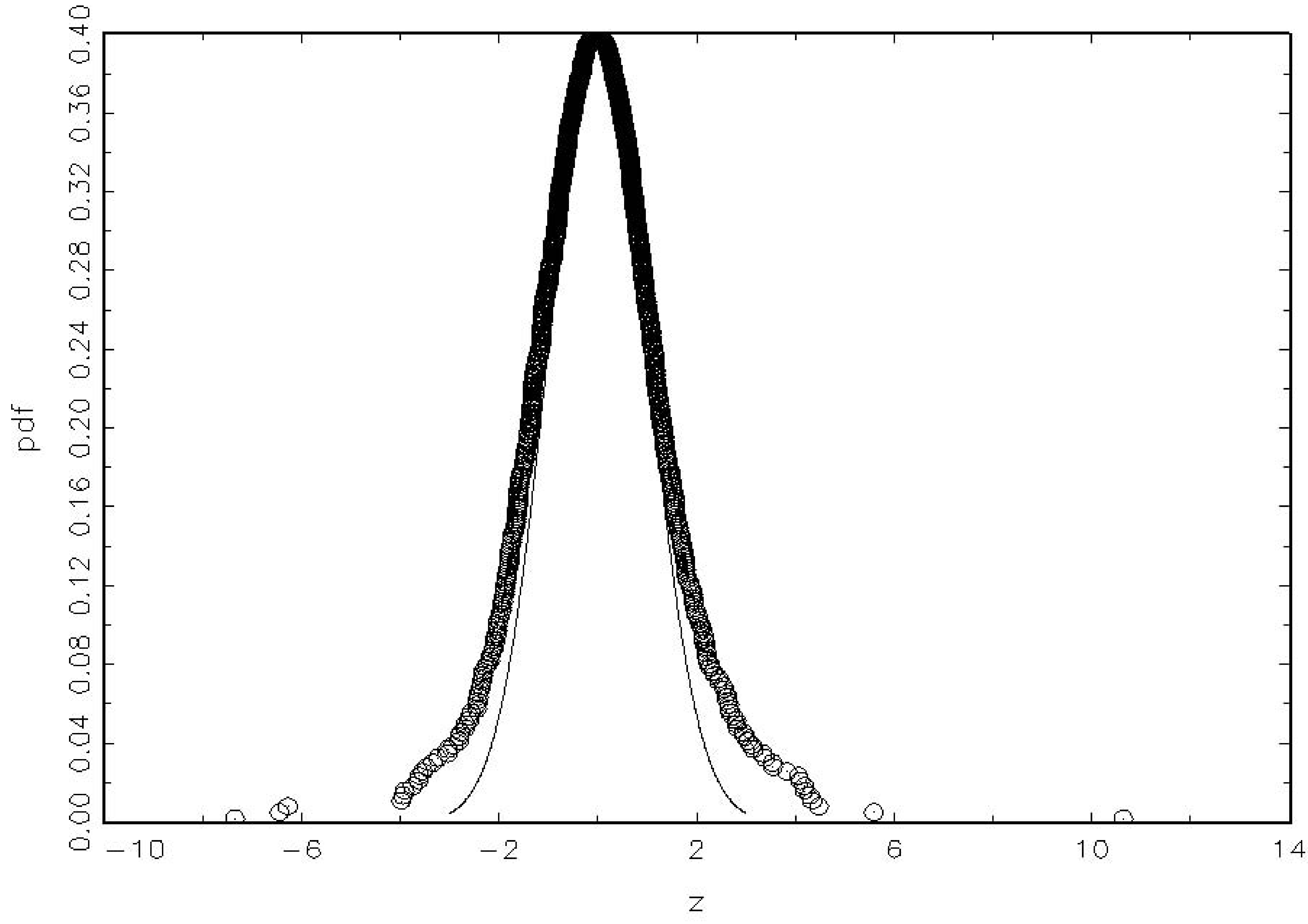
plot $F^{-1}\left(\frac{i-\frac{1}{2}}{q}\right)$ vs. $y_{(i)}$

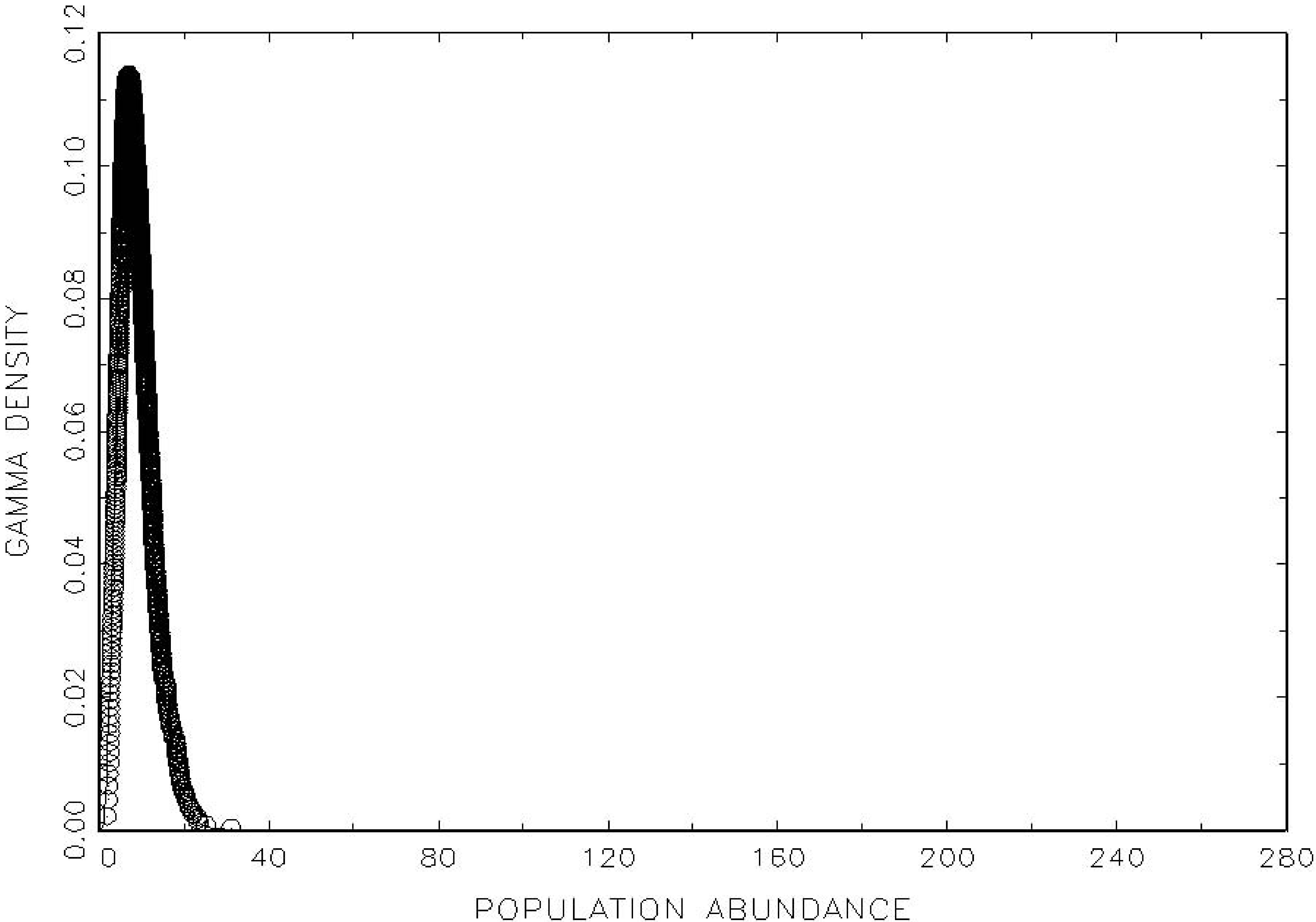
straight line indicates good fit, but the interpretations of departures from fit are not transparent density probability plot (Jones and Daly 1995 , Jones 2004):

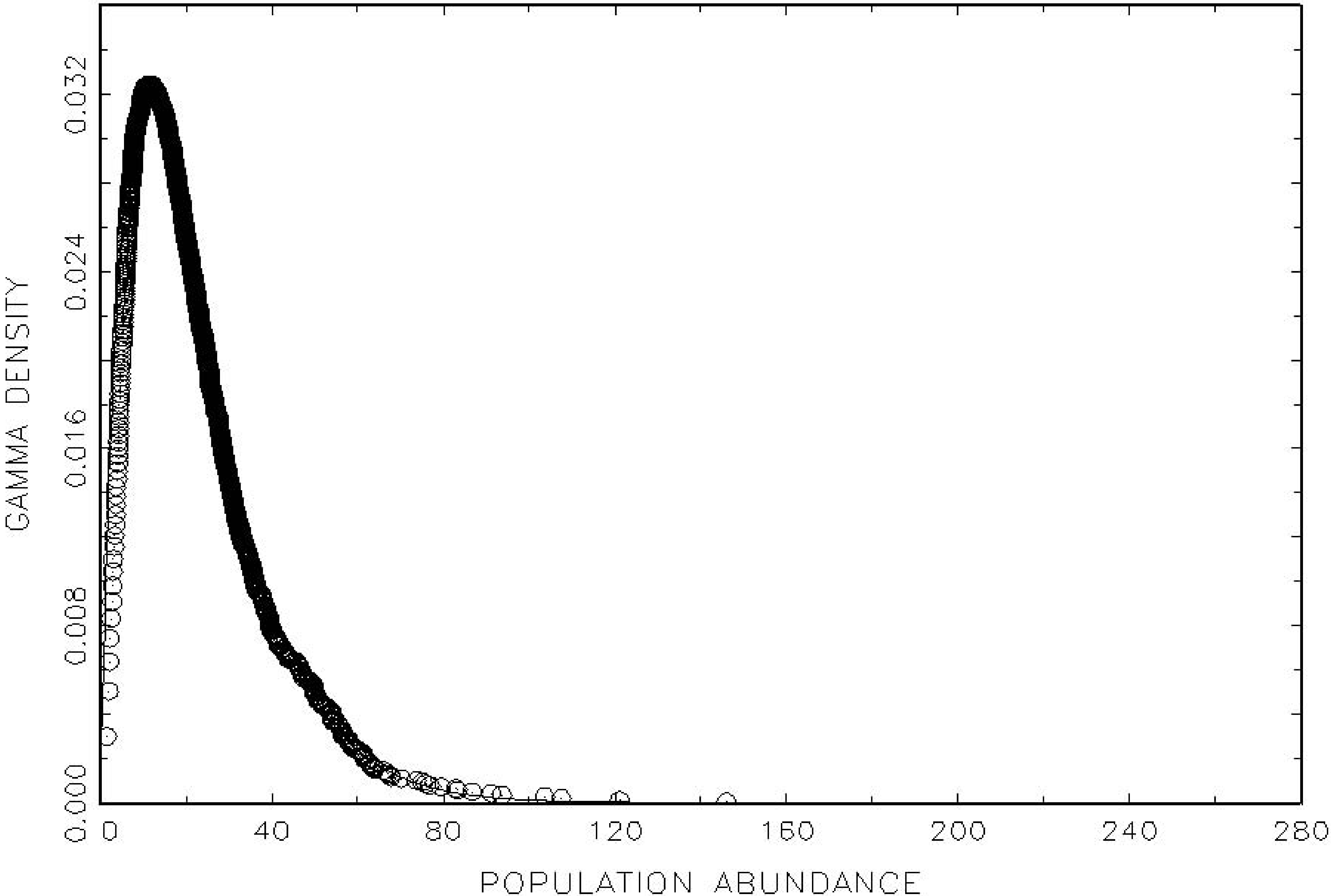
plot $f\left(F^{-1}\left(\frac{i-\frac{1}{2}}{q}\right)\right)$ vs. $y_{(i)}$

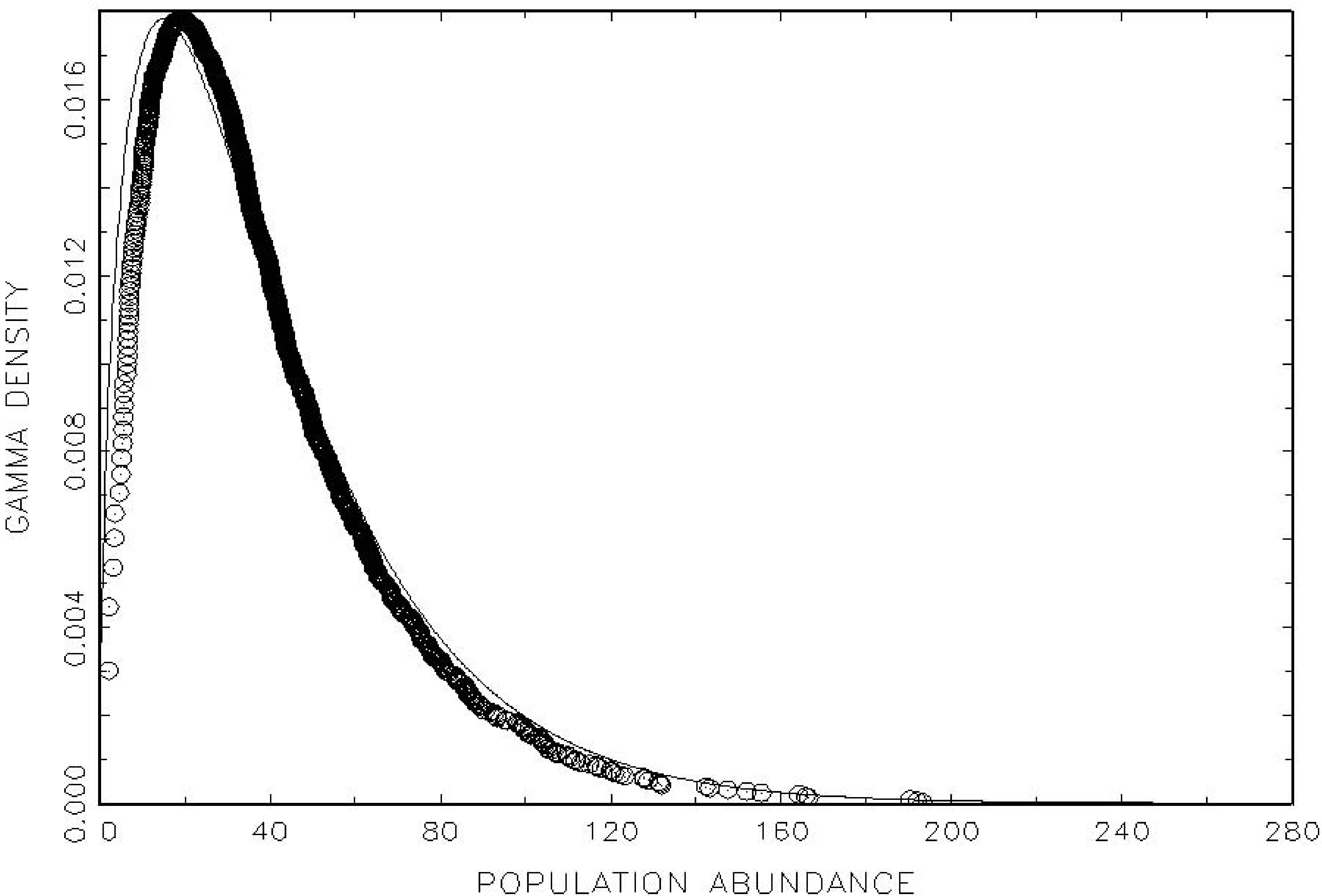
superimposed with plot of $f(y)$


0.00
0.04
0.08
0.12
0.16
0.20
0.24
0.28
0.32
0.36
0.40
pdf
−4
−3
−2
−1
0
1
2
3
4
z

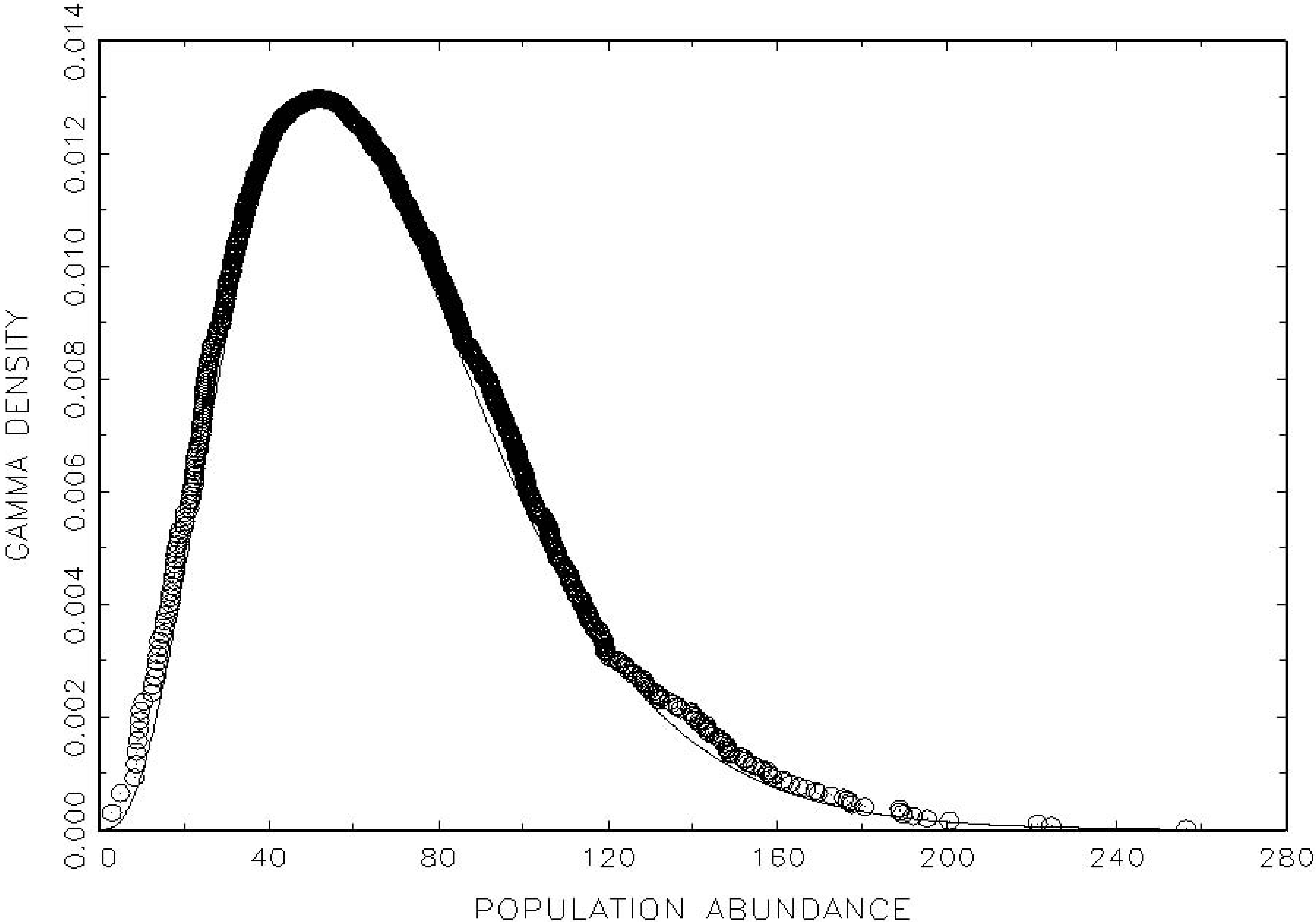

0.00
0.04
0.08
0.12
0.16
0.20
0.24
0.28
0.32
0.36
0.40
pdf
−10
−6
−2
2
6
10
14
z


GAMMA DENSITY
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0
40
80
120
160
200
240
280
POPULATION ABUNDANCE

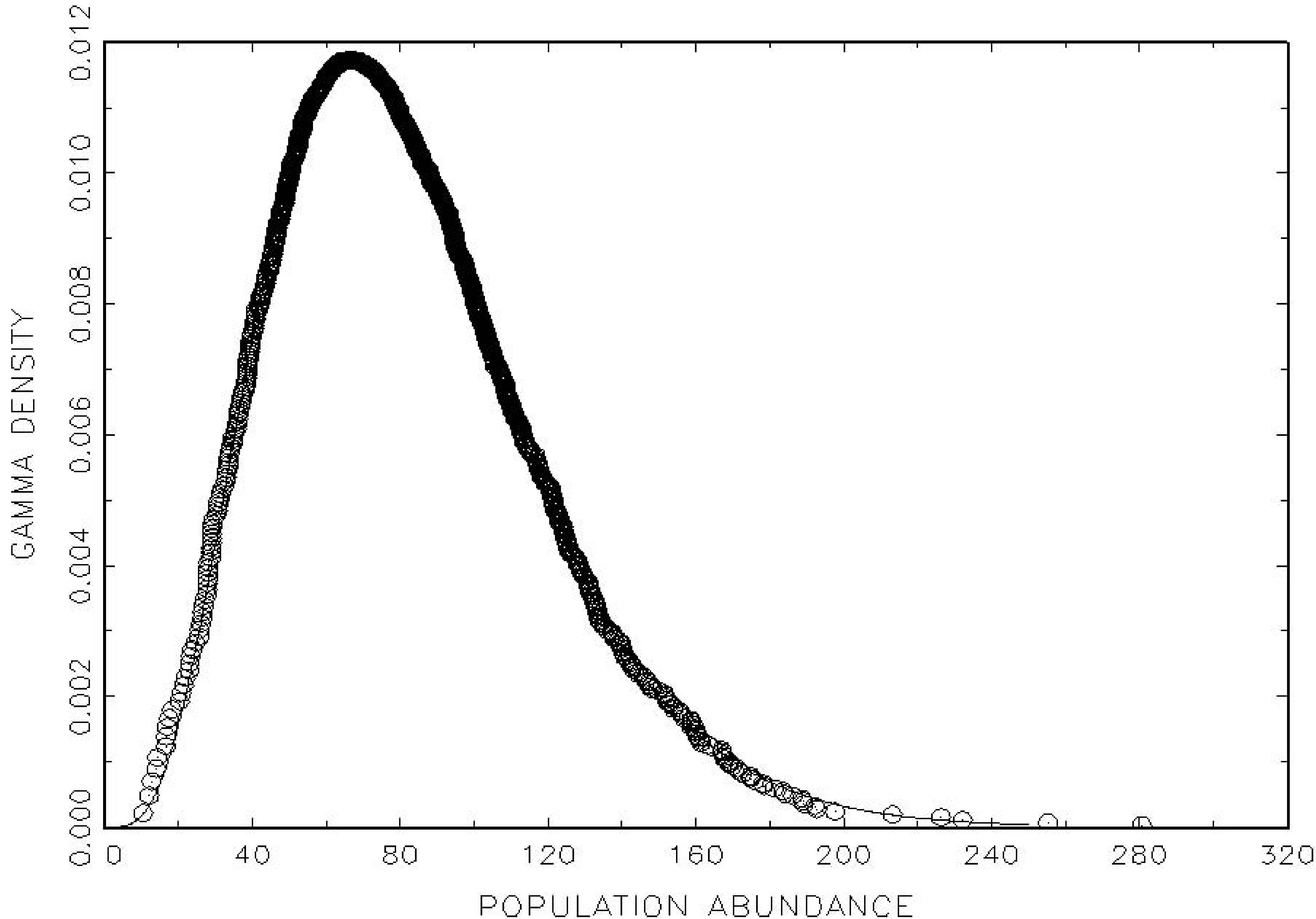

GAMMA DENSITY
0.000
0.008
0.016
0.024
0.032
0
40
80
120
160
200
240
280
POPULATION ABUNDANCE



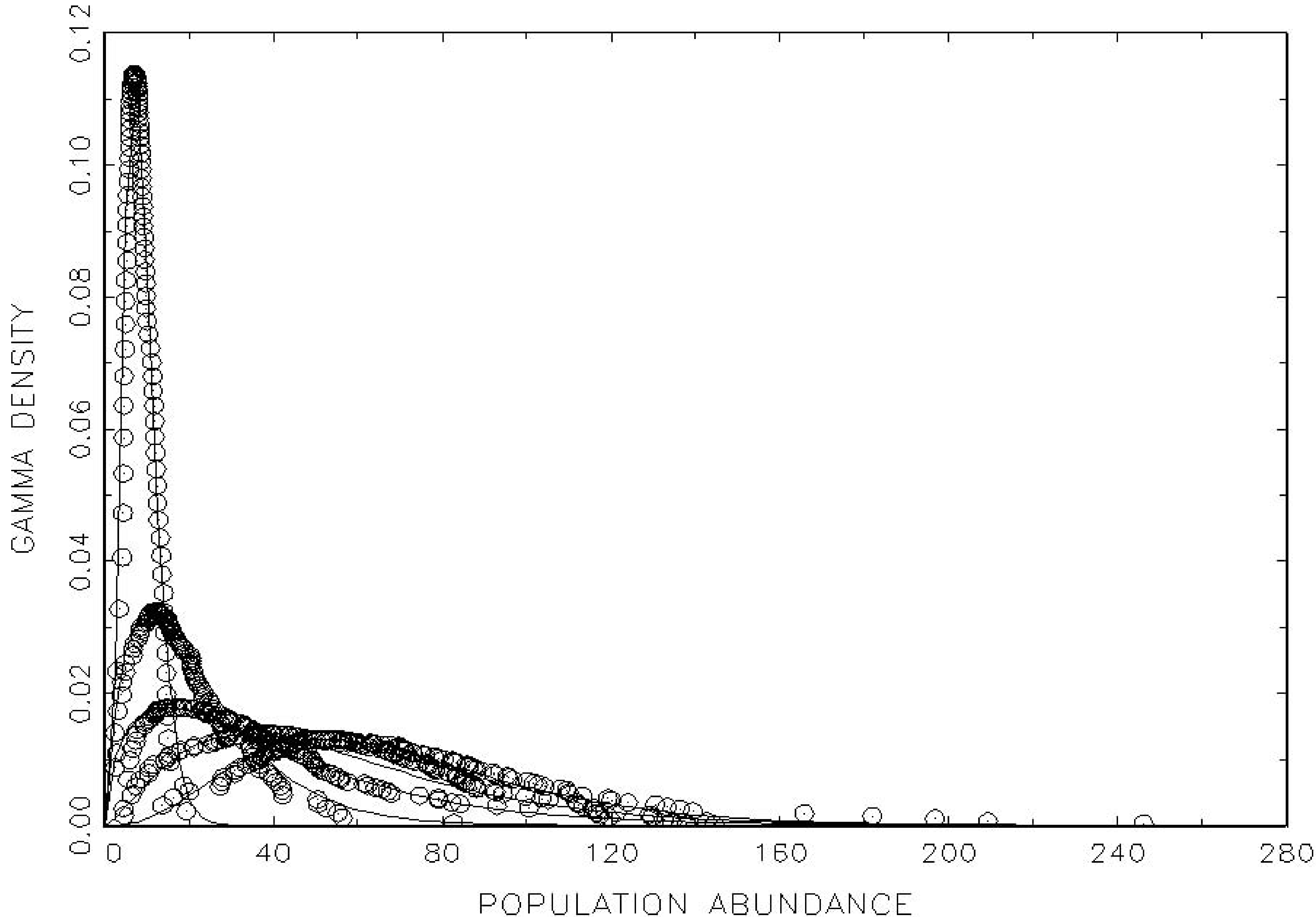
GAMMA DENSITY
0.000
0.004
0.008
0.012
0.016
0
40
80
120
160
200
240
280
POPULATION ABUNDANCE



GAMMA DENSITY
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014
0
40
80
120
160
200
240
280
POPULATION ABUNDANCE



0.000
0.002
0.004
0.006
0.008
0.010
0.012
GAMMA DENSITY
0
40
80
120
160
200
240
280
320
POPULATION ABUNDANCE



GAMMA DENSITY
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0
40
80
120
160
200
240
280
POPULATION ABUNDANCE

Uses of the stochastic logistic: data analysis!

r, k, σ^2 unknown parameters

0	t_1	t_2	$\cdots$	t_q
n_0	n_1	n_2	$\cdots$	n_q

time series data (pop. abundances)

$\tau_1 = t_1 - 0$, $\tau_2 = t_2 - t_1$, ..., $\tau_q = t_q - t_{q-1}$ intervals

Likelihood function:

$$L(r, k, \sigma^2) \;=\; p(n_1, \tau_1 \,|\, n_0)p(n_2, \tau_2 \,|\, n_1)\cdots p(n_q, \tau_q \,|\, n_{q-1})$$

$\widehat{r}$, $\widehat{k}$, $\widehat{\sigma}^2$ maximum likelihood estimates (values of r, k, σ^2 that jointly maximize $L(r, k, \sigma^2)$)

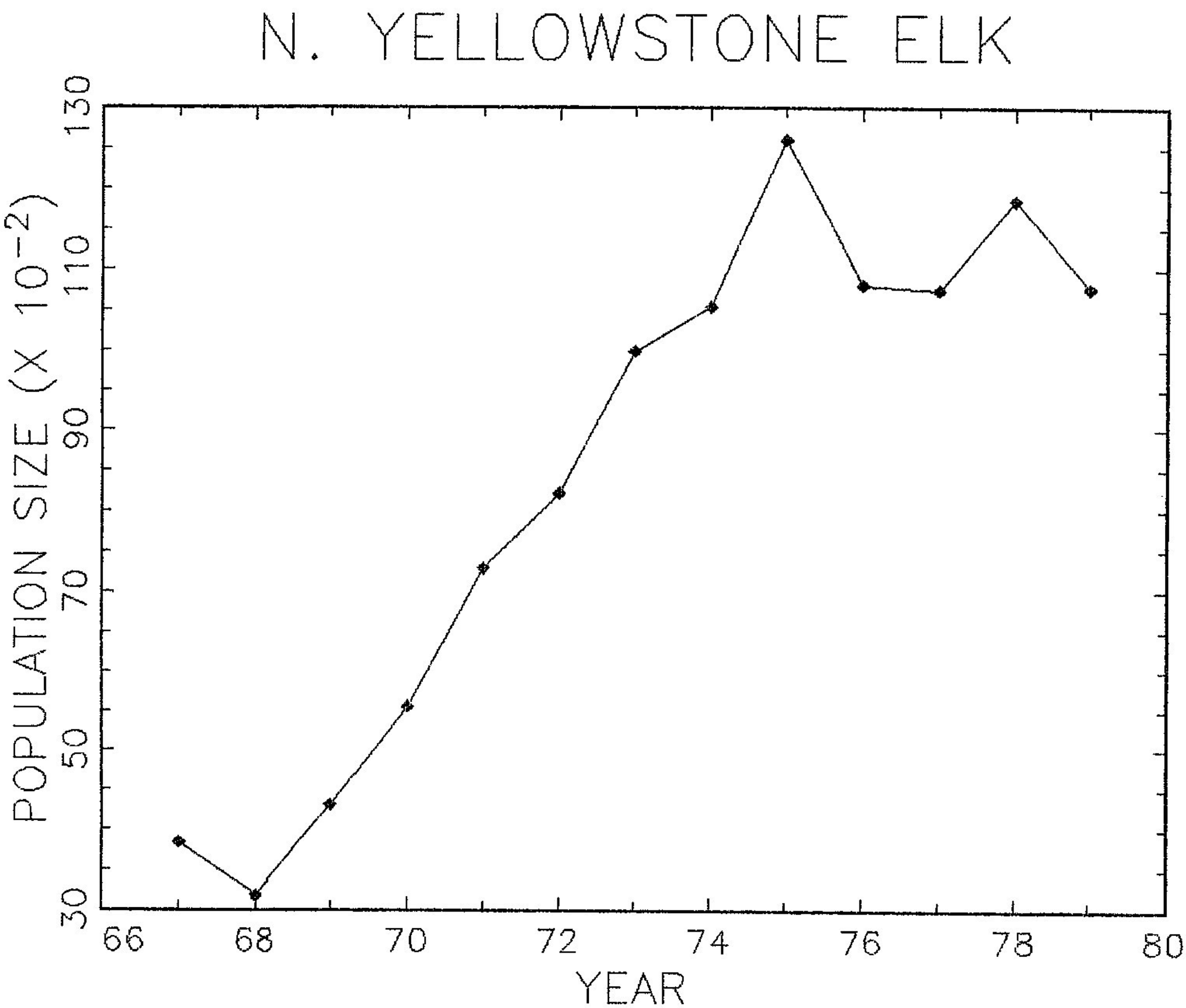
requires numerical maximization: optim() function in R, etc.

Residuals:

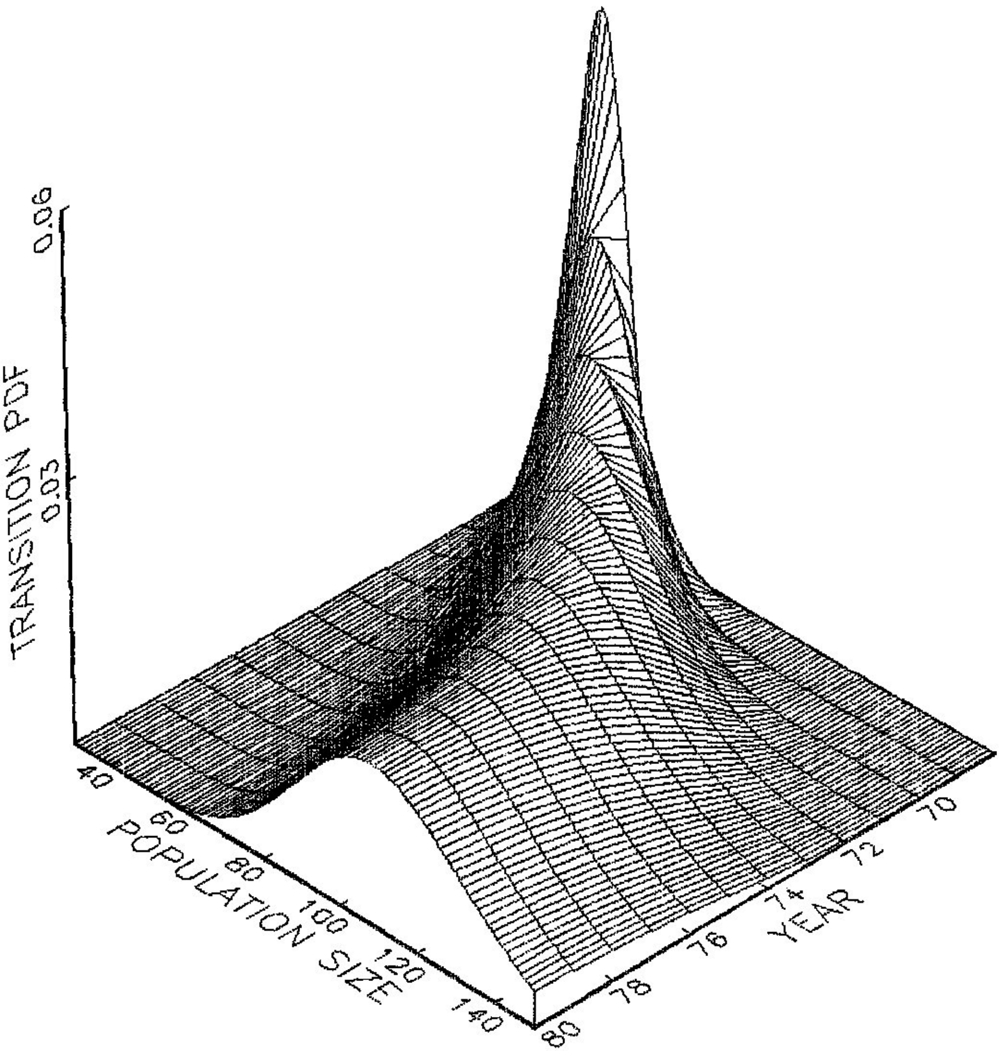
$$z_i = \frac{\left(\frac{\widehat{\beta}_i n_i}{\widehat{\alpha}_i}\right)^{1/3} - \left(1 - \frac{1}{9\widehat{\alpha}_i}\right)}{\sqrt{9\widehat{\alpha}_i}}$$

are approximately normal(0, 1) distributed, where

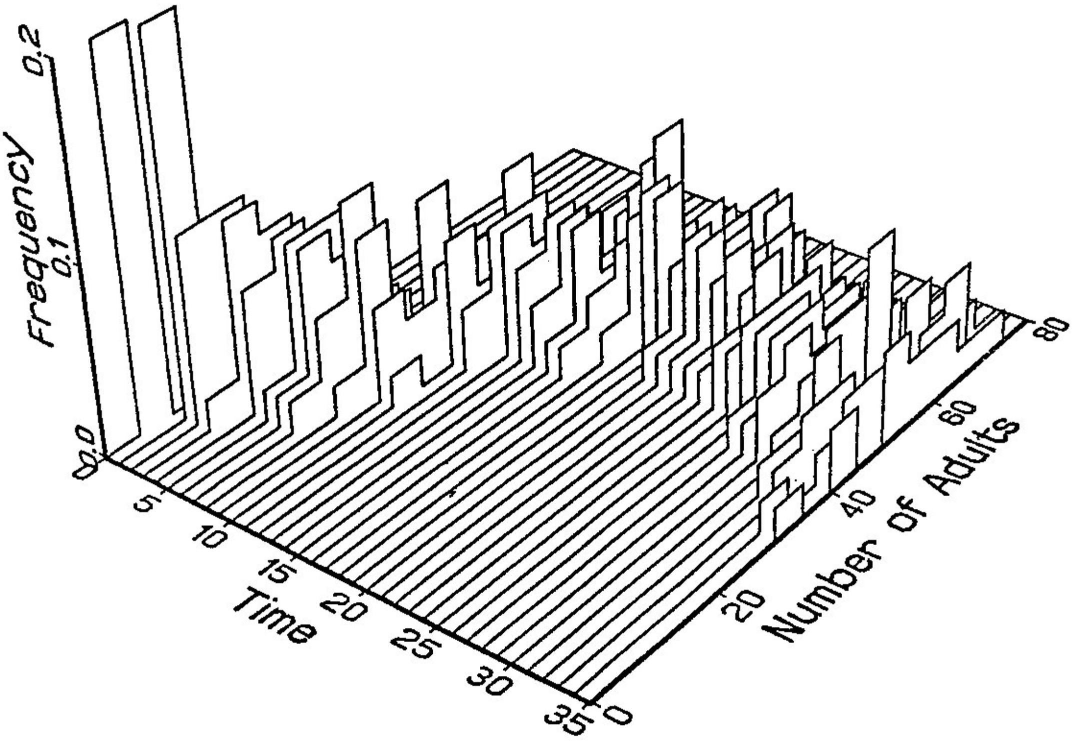
$$\widehat{\alpha}_i = \alpha_{\tau_i}\left(n_{i-1}, \widehat{r}, \widehat{k}, \widehat{\sigma}^2\right)$$
$$\widehat{\beta}_i = \beta_{\tau_i}\left(n_{i-1}, \widehat{r}, \widehat{k}, \widehat{\sigma}^2\right)$$



N. YELLOWSTONE ELK
POPULATION SIZE (X 10−2)
30
50
70
90
110
130
66
68
70
72
74
76
78
80
YEAR

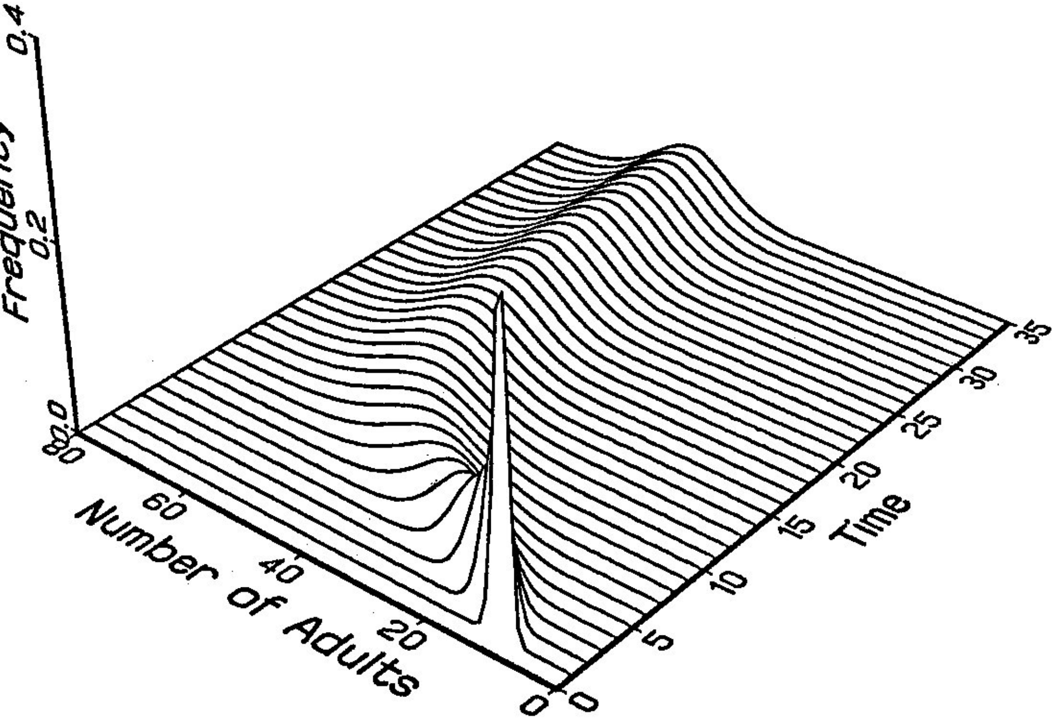

TRANSITION PDF
0.06
0.03
40
60
80
100
120
140
POPULATION SIZE
80
78
76
74
72
70
YEAR

BRAZIL



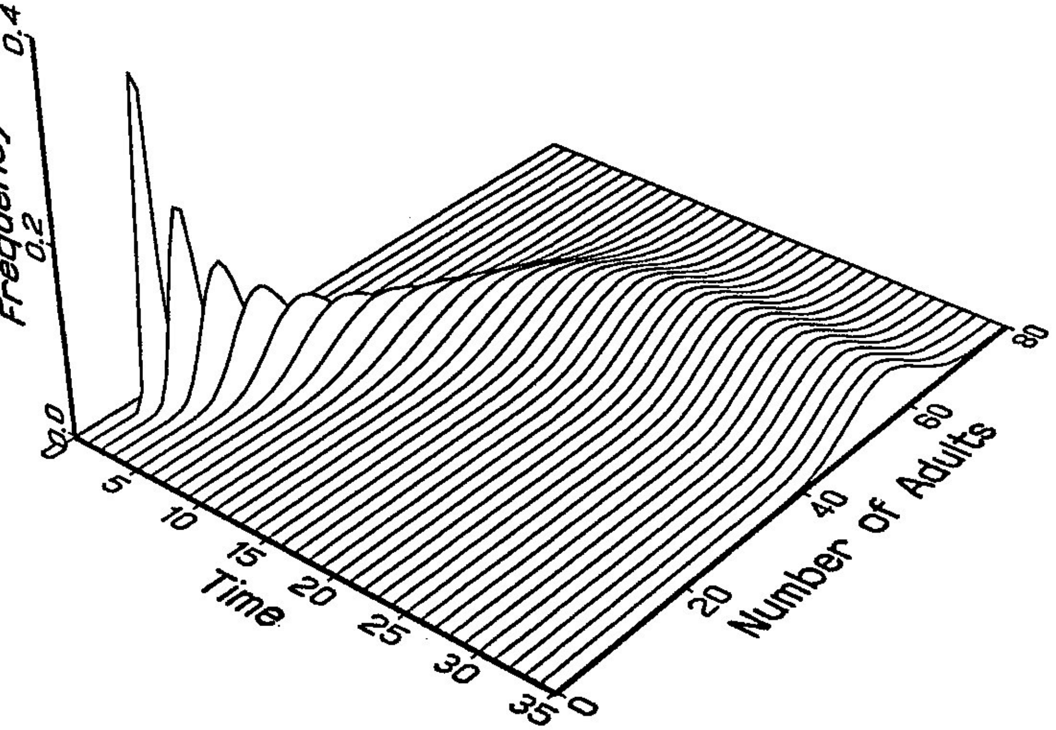
Frequency
0.0
0.1
0.2
Time
5
10
15
20
25
30
35
Number of Adults
0
20
40
60
80

(a)

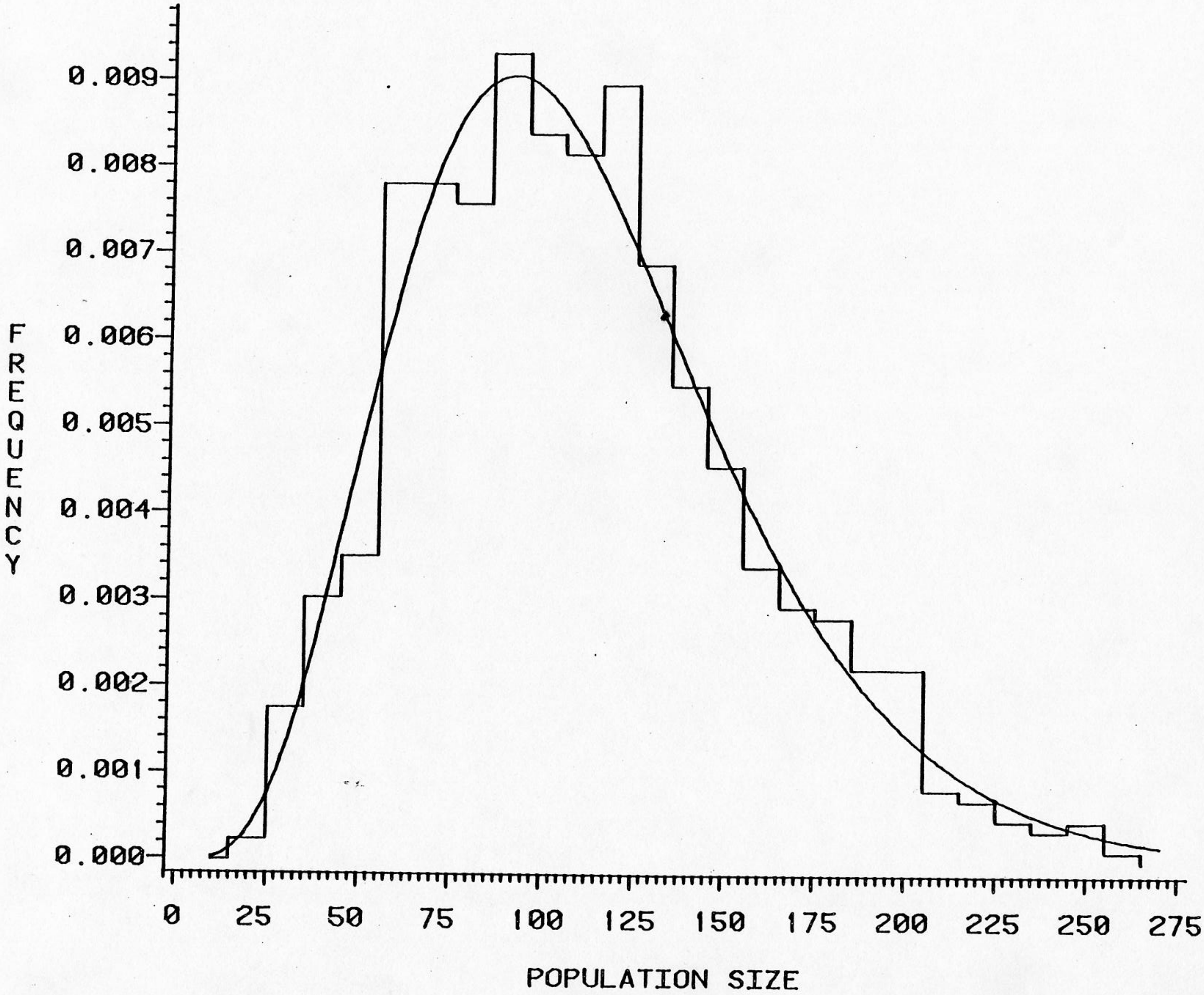


Frequency
0.4
0.2
0.0
80
60
40
20
0
Number of Adults
0
5
10
15
20
25
30
35
Time

(b)



Frequency
0.4
0.2
0.0
0
5
10
15
20
25
30
35
Time
0
20
40
60
80
Number of Adults



FREQUENCY
0.009
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000
0 25 50 75 100 125 150 175 200 225 250 275
POPULATION SIZE

brian@uidaho.edu

References: Chapman 1928 *Ecology*, Dennis 1978 *Mathematical Biosciences,* Dennis 1989a *Natural Resource Modeling*, Dennis 1989b in MacDonald et al. eds. *Estimation and analysis of insect populations*, Dennis and Patil 1984 *Mathematical Biosciences*, Hutchinson 1978 *An introduction to population ecology*, Jones 2004 *The American Statistician*, Jones and Daly 1995 *Communications in Statistics—Simulation and Computation*, Leach 1981 *JRSS* A, Lotka 1925 Elements of physical biology, Pearl and Reed 1920 *PNAS*, Reed and Berkson 1929 *J Physical Chemistry*, Verhulst 1838 Correspondance Mathématique et Physique, Wieszak 1988 UI PhD Thesis, Williams 1972 in Deevey, ed. *Growth by intussusception: ecological essays in honor of G. E. Hutchinson*.