Modeling population size

Imagine a biological population; yeast cells, for example.
Start with a few in some sugar water.
They begin reproducing. Assume no death.
Can we model the population size as a function of time?
Usual model in calculus:

Population size m_t satisfies $\frac{dm_t}{dt} = \beta m_t$,

$m_t = m_0 e^{\beta t}$ exponential growth

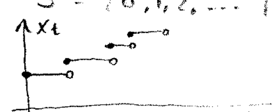
This is a deterministic model.
It cannot be the whole story.
How long does it take for the population size to double?

$$m_t = 2m_0 = m_0 e^{\beta t}$$
$$2 = e^{\beta t}$$
$$t = \frac{1}{\beta} \ln(2),$$ always the same!

Stochastic process model:

X_t is population size at time t, values in $S = \{0, 1, 2, \dots\}$.

X increases over time by jumps of size 1.



Is it enough to know the total population size at time t in order to assess the probabilities of various future events?

Not really!
One would like to know the age at which cells reproduce, how "young" or "old" the population is, etc.
This is too much!

If we begin with cells of varying ages, and plenty of them, then a Markov model may be reasonable.

Imagine 1000 cells of various ages.
The number of offspring they produce in $[0,1]$ tells us very little about the number in $[1,2]$, etc.
A Poisson model is appropriate.

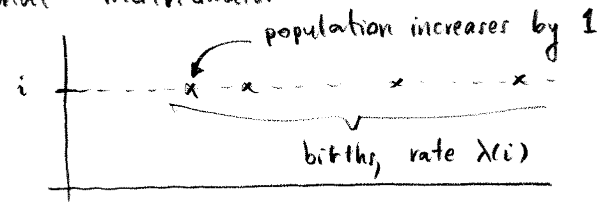
Similar to the arrivals of phone calls, as we discussed earlier.

Suppose we model the population size X_t as a Markov process.

We need to determine the sojourn parameters $\lambda(i)$ and jump matrix Q.

Q is the easiest: $Q(i,i+1) = 1$, for $i = 1, 2, \ldots$

The parameter $\lambda(i)$ is the rate at which a population of i individuals gives birth to additional individuals.



$\lambda(i)$ will be proportional to i.

The more individuals, the higher the rate at which new individuals will be born.

$$\lambda(i) = \beta i$$

β is the rate at which <u>one</u> individual gives birth (# births per hour

Yule process — the name for this pure birth process

$$\lambda(i) = \beta \cdot i$$

$$\lambda(0) = 0 \qquad (\text{0 population, 0 rate of births})$$

$$Q(i,j) = \begin{cases} 1 & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases}$$

Simulations yule.m

X_0 is the initial population size. Random for a packet of baking yeast.

Generate times $T_1, T_2, T_3, \ldots$ of jumps.

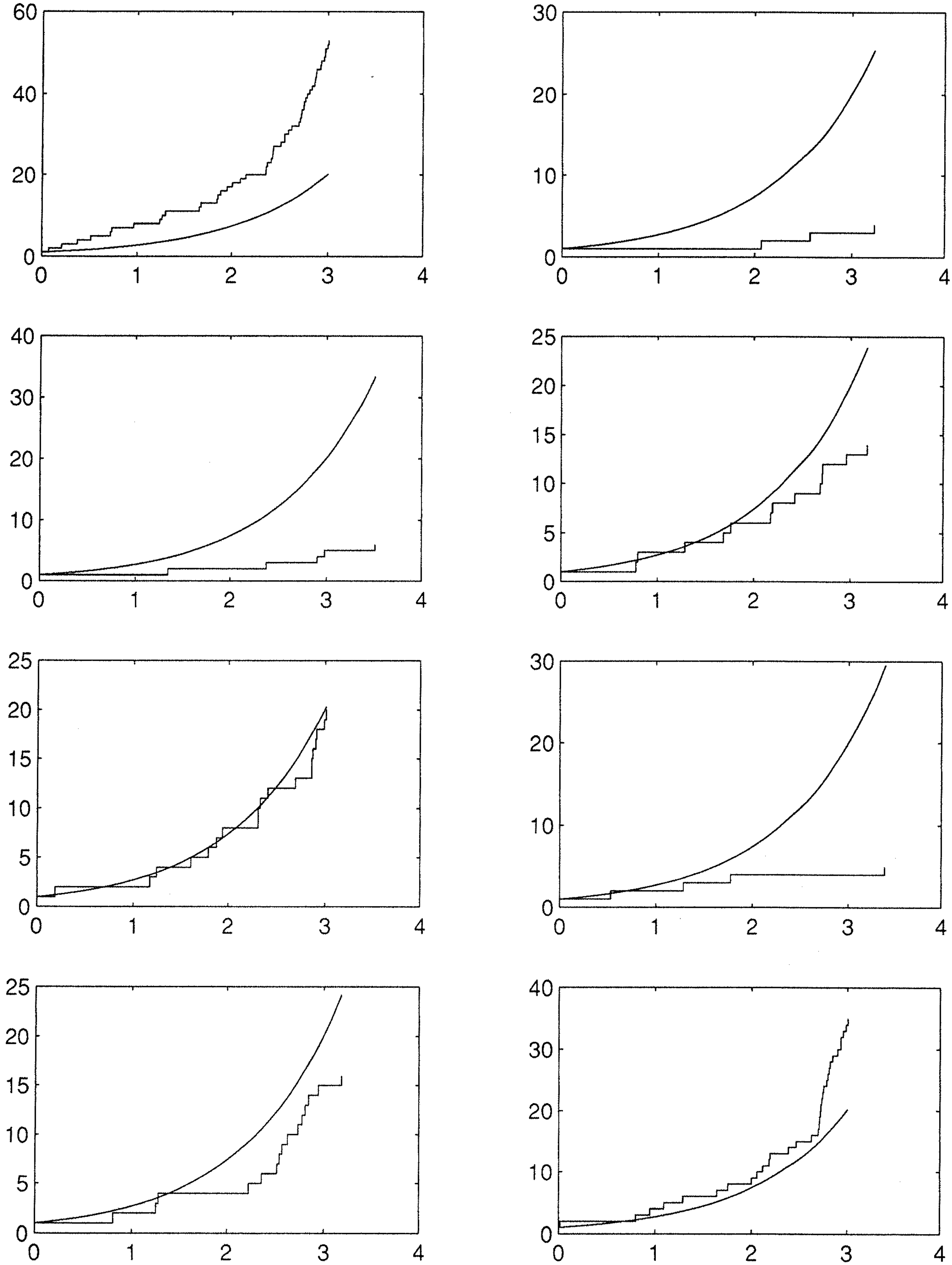
$T_1 =$ exponential, mean $\frac{1}{\beta \cdot X_0}$

$T_2 = T_1 +$ exponential, mean $\frac{1}{\beta(X_0+1)}$

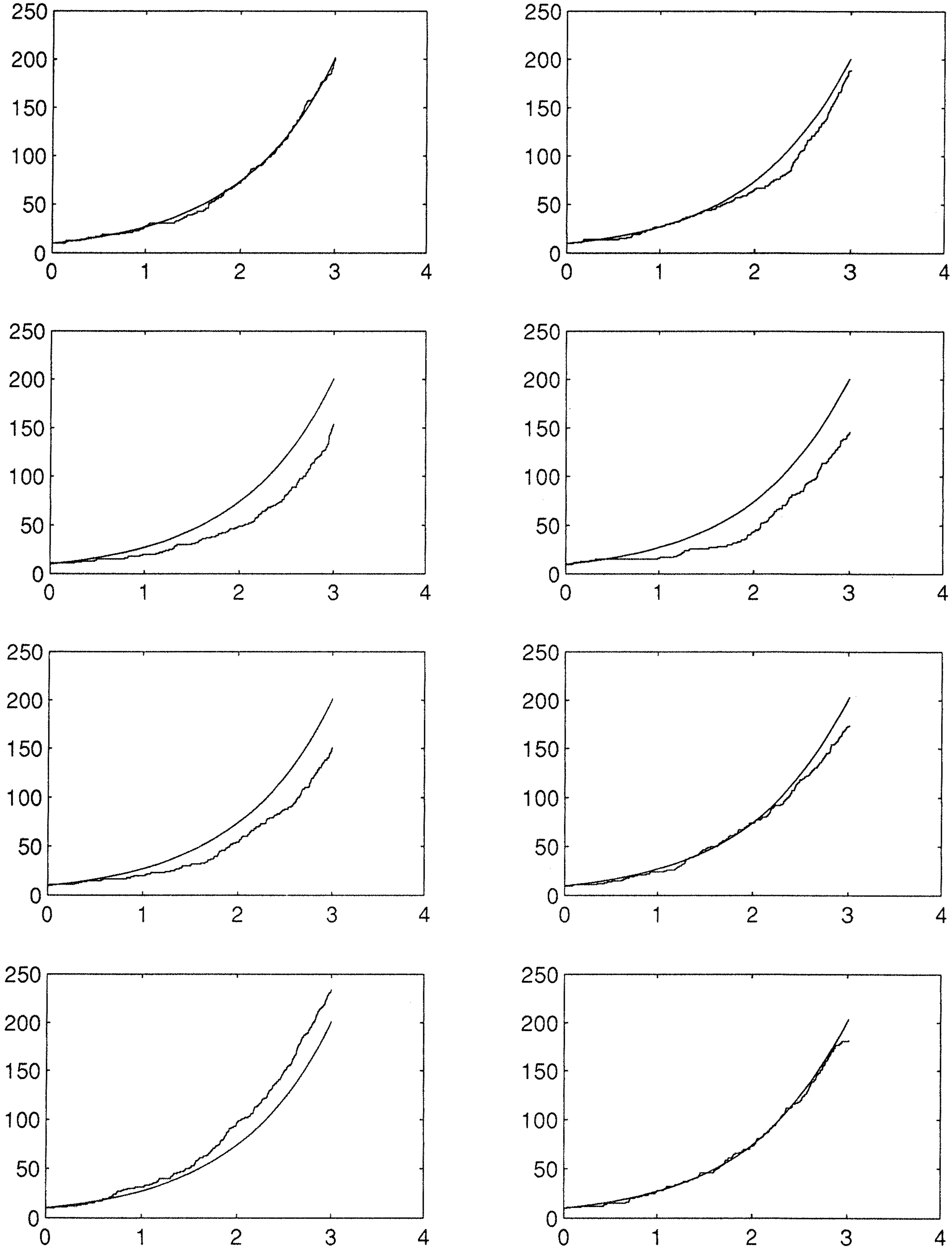
$\vdots$

$T_{n+1} = T_n +$ exponential, mean $\frac{1}{\beta(X_0+n)}$

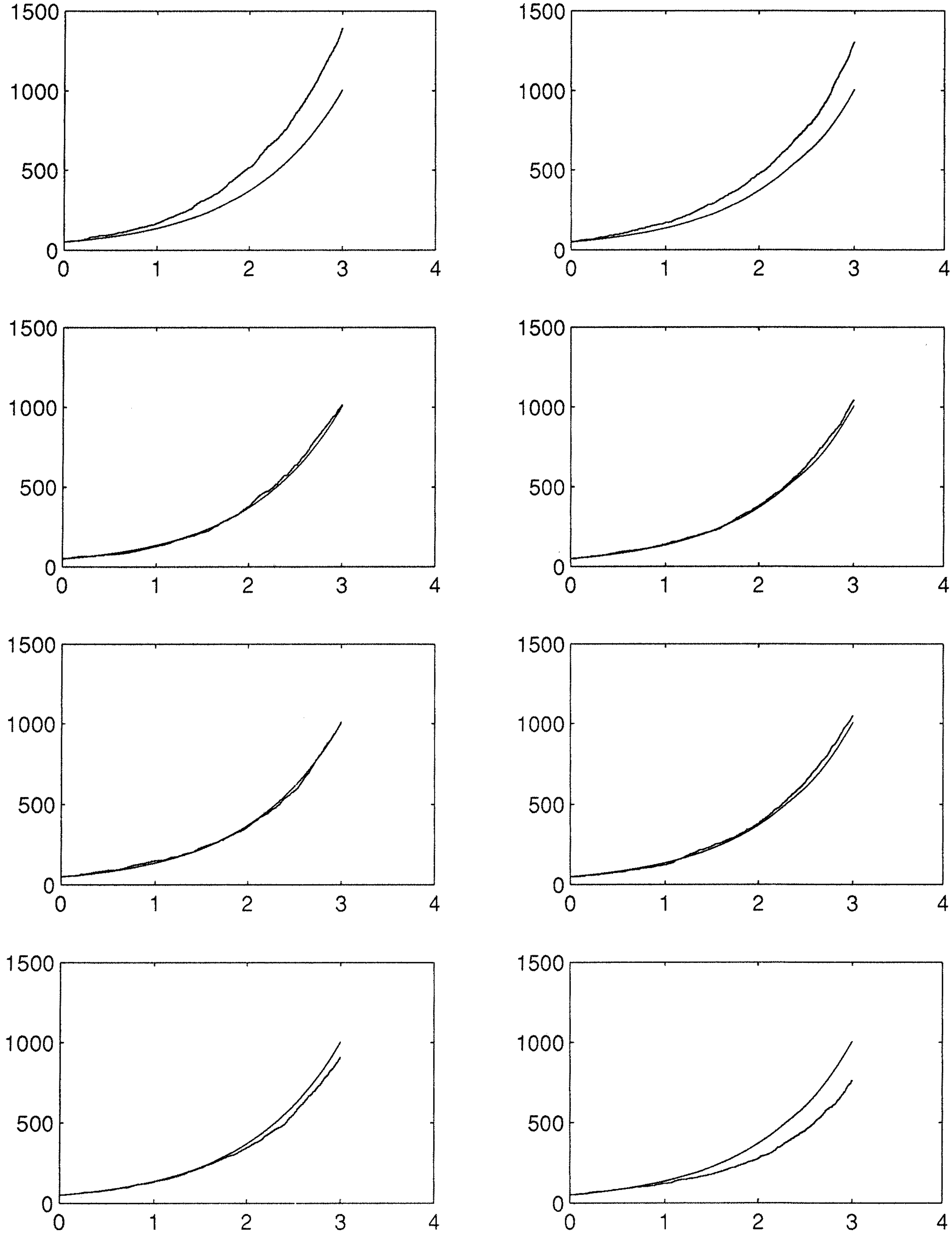
Yule process, $X_0=1$, $\beta=1$



Yule process, $X_0 = 10$, $\beta = 1$



Yule process, $X_0 = 50$, $\beta = 1$.



Transition probabilities

$$P_{ij}(t) = \mathbb{P}(X_t = j \mid X_0 = i)$$

- Of course, $P_{ij}(t) = 0$ if $j < i$.

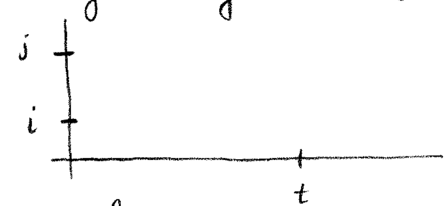
- $$\begin{aligned} P_{ii}(t) &= \mathbb{P}(X_t = i \mid X_0 = i) \\ &= \mathbb{P}(\text{no births in } [0,t] \mid X_0 = i) \\ &= \mathbb{P}(T_1 > t \mid X_0 = i) \\ &= e^{-\lambda(i)\cdot t} \end{aligned}$$

  Notice that $\frac{d}{dt} P_{ii}(t) = -\lambda(i)\, P_{ii}(t)$.

- $P_{ij}(t)$ for $j > i$ is harder to find.

  It is easiest to begin with a differential equation for $P_{ij}(t)$.

$$\frac{d}{dt} P_{ij}(t) = \text{rate of change of probability of being in state } j$$



$$= P_{i,j-1}(t) \cdot \lambda(j-1) \quad \text{rate of jumping into state } j$$

$$- P_{i,j}(t) \cdot \lambda(j) \quad \text{rate of leaving state } j$$

$$= \beta(j-1)\, P_{i,j-1}(t) - \beta \cdot j\, P_{i,j}(t)$$

(We will do this more rigorously later.)

Solution: (Taylor & Karlin, p. 339)

$$P_{1,n}(t) = e^{-\beta t}(1 - e^{-\beta t})^{n-1}$$

It is very rare that exact solutions are available!

Mean population size

$$
\begin{aligned}
m_t &= \mathbb{E}[X_t \mid X_0 = i] \\
&= \sum_{j=i}^{\infty} j \cdot \mathbb{P}(X_t = j \mid X_0 = i) \\
&= \sum_{j=i}^{\infty} j \cdot P_{ij}(t)
\end{aligned}
$$

$$
\begin{aligned}
\frac{d}{dt} m_t &= \sum_{j=i}^{\infty} j \cdot \frac{d}{dt} P_{ij}(t) \\
&= \sum_{j=i}^{\infty} j \left[\beta \cdot (j-1) \cdot P_{i,j-1}(t) - \beta \cdot j \cdot P_{ij}(t) \right] \\
&= i \cdot \beta \cdot \overset{\to\, 0,\ \text{no death}}{\cancel{(i-1) \cdot P_{i,i-1}(t)}} + \sum_{j=i+1}^{\infty} \beta \cdot j(j-1)\, P_{i,j-1}(t) \\
&\qquad + \sum_{j=i}^{\infty} -\beta \cdot j^2 \cdot P_{ij}(t) \\
&= \sum_{j=i}^{\infty} \beta \cdot (j+1)(j)\, P_{ij}(t) - \sum_{j=i}^{\infty} \beta \cdot j^2 \cdot P_{ij}(t) \\
&= \sum_{j=i}^{\infty} \beta \cdot j \cdot P_{ij}(t) \\
&= \beta \cdot m_t,
\end{aligned}
$$

So the mean population size follows exactly the usual exponential growth model!

$$m_t = i\, e^{\beta t} \qquad \text{when} \quad X_0 = i.$$

- The simulations plot outcomes of population size versus the mean. They follow the mean pretty well when i is large.

- Let $V_t = \mathrm{Var}(X_t)$.
  In much the same way you can show that $\frac{d}{dt} V_t = \beta \cdot (2V_t + m_t)$, and then find that

$$
\begin{aligned}
V_t &= i\,(e^{2\beta t} - e^{\beta t}) \\
SD_t &= \sqrt{V_t} \\
\frac{SD_t}{m_t} &= \frac{1}{\sqrt{i}} \cdot \sqrt{1 - e^{-\beta t}},
\end{aligned}
$$

so this is small as $i \to \infty$.

Computing transition probabilities

So far, what we know about Markov processes with a countable state space S is that they stay in state i for an exponentially distributed length of time with mean $\frac{1}{\lambda(i)}$, then jump to state j with probability Q_{ij}. This allows us to simulate outcomes of Markov processes.

What we know about the transition matrices $P(t)$, $t \geq 0$ is that:

- $P_{ij}(t) = \mathbb{P}(X_t = j \mid X_0 = i)$ (the definition)
- $P(t+s) = P(t)P(s)$ (the Chapman-Kolmogorov eqn.)

However, we don't know how to compute $P(t)$ for any particular value of t (except $t=0$, of course).

If we could compute $P(1)$, say, then we would have $P(2) = P(1)P(1) = P(1)^2$, $P(3) = P(1)^3$, etc.
We could look at $\lim_{n\to\infty} P(1)^n$ and talk about limiting distributions.
If X had an initial distribution μ, then X_1 would have distribution $\mu P(1)$, X_2 would have distribution $\mu P(1)^2$, and so on, and we could talk about the invariant distribution.

What we need is a way to get started, to compute $P(t)$ for some non-zero value of t.

We start with an important lemma, then derive a differential equation for $P(t)$, then show how to solve it.

Computing transition probabilities

Lemma For all i, j in S and $t > 0$,

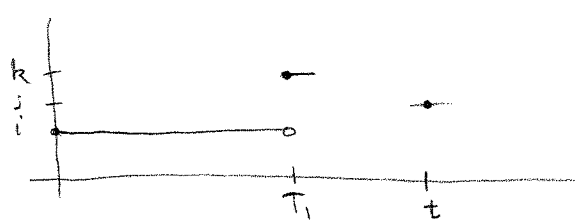
$$P_{ij}(t) = e^{-\lambda(i)t} \cdot \delta_{ij} + \int_0^t \lambda(i) e^{-\lambda(i) \cdot s} \sum_{k \in S} Q_{ik} P_{kj}(t-s)\, ds$$

Proof If i is absorbing, then $\lambda(i) = 0$, $P_{ij}(t) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i. \end{cases}$

If i is stable, then $0 < \lambda(i) < \infty$, and the time T_1 of the first jump is finite.

$$\begin{aligned} P_{ij}(t) &= \mathbb{P}(X_t = j \mid X_0 = i) \\ &= \mathbb{P}(X_t = j \text{ and } T_1 > t \mid X_0 = i) + \mathbb{P}(X_t = j \text{ and } T_1 \le t \mid X_0 = i) \\ &= \quad \text{①} \quad + \quad \text{②} \end{aligned}$$

$$\begin{aligned} \text{①} &= \begin{cases} 0 & \text{if } j \neq i \\ \mathbb{P}(T_1 > t \mid X_0 = i) & \text{if } j = i \end{cases} \\ &= \begin{cases} 0 & \text{if } j \neq i \\ e^{-\lambda(i)t} & \text{if } j = i \end{cases} \\ &= e^{-\lambda(i)t}\, \delta_{ij} \end{aligned}$$



$$\begin{aligned} \text{②} &= \sum_{k \in S} \mathbb{P}(X_t = j,\, X_{T_1} = k,\, T_1 \le t \mid X_0 = i) \\ &= \sum_{k \in S} \mathbb{P}(X_t = j,\, T_1 \le t \mid X_{T_1} = k, X_0 = i) \cdot \mathbb{P}(X_{T_1} = k \mid X_0 = i) \\ &= \sum_{k \in S} \mathbb{E}\, \mathbb{P}(X_t = j, T_1 \le t \mid X_{T_1} = k, T_1, X_0 = i) \cdot Q_{ik} \\ &= \sum_{k \in S} \mathbb{E}\, 1_{\{T_1 \le t\}}\, \mathbb{P}(X_t = j \mid X_{T_1} = k, T_1, X_0 = i)\, Q_{ik} \\ &= \sum_{k \in S} \mathbb{E}\, 1_{\{T_1 \le t\}}\, P_{kj}(t - T_1)\, Q_{ik} \\ &= \sum_{k \in S} Q_{ik} \cdot \int_0^\infty ds\, 1_{\{s \le t\}}\, P_{kj}(t-s) \underbrace{\lambda(i) e^{-\lambda(i)s}}_{\text{density of } T_1} , \end{aligned}$$

strong Markov and some conditioning

and we are done.

<u>Lemma</u> For all i and j in S and $t > 0$,

$$P_{ij}(t) = e^{-\lambda(i)t}\delta_{ij} + \sum_{k\in S}\int_0^t \lambda(i)e^{-\lambda(i)r} Q_{ik} P_{kj}(t-r)\,dr$$

<u>Proof</u> If i is absorbing, then $\lambda(i) = 0$ and $P_{ij}(t) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i, \end{cases}$

so the claim is true in this case.

If i is not absorbing, then $0 < \lambda(i) < \infty$ and, starting in state i, the time T_1 of the first jump is finite.

$$\begin{aligned} P_{ij}(t) &= \mathbb{P}(X_t = j \mid X_0 = i) \\ &= \mathbb{P}(X_t = j \text{ and } T_1 > t \mid X_0 = i) + \mathbb{P}(X_t = j \text{ and } T_1 \le t \mid X_0 = i) \\ &= \quad ① \quad + \quad ② \end{aligned}$$

$$① = \begin{cases} 0 & \text{if } j \neq i \\ \mathbb{P}(T_1 > t \mid X_0 = i) & \text{if } j = i \end{cases}$$

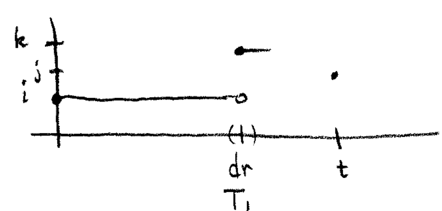
$$= \begin{cases} 0 & \text{if } j \neq i \\ e^{-\lambda(i)t} & \text{if } j = i \end{cases}$$

$$= e^{-\lambda(i)t}\delta_{ij}.$$

$$② = \sum_{k\in S} \mathbb{P}(X_t = j, X_{T_1} = k, T_1 \le t \mid X_0 = i)$$ law of total prob.

$$= \sum_{k\in S}\int_0^t \mathbb{P}(X_t = j, X_{T_1} = k, T_1 \in dr \mid X_0 = i)$$ law of total prob. and some lack of rigor

$$= \sum_{k\in S}\int_0^t \mathbb{P}(X_t = j \mid X_{T_1} = k, T_1 \in dr, X_0 = i)\,\mathbb{P}(X_{T_1} = k \mid T_1 \in dr, X_0 = i)\,\mathbb{P}(T_1 \in dr \mid X_0 = i)$$



picture supporting using the strong Markov property at time T_1

$$= \sum_{k\in S}\int_0^t \underbrace{P_{kj}(t-r)}_{\text{go from } k \text{ to } j \text{ over time } t-r} \cdot \underbrace{Q_{ik}}_{\text{jump prob.}} \cdot \underbrace{\lambda(i)e^{-\lambda(i)r}\,dr}_{\text{density of } T_1}$$

Rearranging completes the proof.

Differential equation for $P(t)$

First, rewrite the formula from the lemma by changing variables:

$$P_{ij}(t) = e^{-\lambda(i)t}\delta_{ij} + \sum_{k\in S}\int_0^t \lambda(i)\, e^{-\lambda(i)r} Q_{ik} P_{kj}(t-r)\,dr$$

$$\left[\begin{array}{ll} u = t-r & r=0,\ u=t \\ du = -dr & r=t,\ u=0 \end{array}\right]$$

$$= e^{-\lambda(i)t}\delta_{ij} + \sum_{k\in S}\int_0^t \lambda(i)\, e^{-\lambda(i)(t-u)} Q_{ik} P_{kj}(u)\,du$$

$$= e^{-\lambda(i)t}\left[\delta_{ij} + \sum_{k\in S}\int_0^t \lambda(i)\, e^{\lambda(i)u} Q_{ik} P_{kj}(u)\,du\right]$$

Technical note: We assumed earlier that $\lim_{t\downarrow 0} P(t) = I$, from which we showed that $P_{ij}(t)$ is continuous as a function of t. Thus, the integrand here is continuous. Thus, the left side, $P_{ij}(t)$, is differentiable in t. (Repeating this, it is infinitely differentiable!)

Differentiate with respect to t:

$$\frac{d}{dt}P_{ij}(t) = -\lambda(i)\cdot\left[\qquad P_{ij}(t) \qquad\right]$$

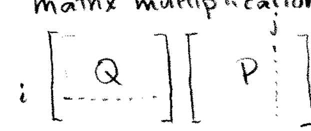
$$+ e^{-\lambda(i)t}\cdot\left[0 + \sum_{k\in S}\lambda(i)\, e^{\lambda(i)t} Q_{ik} P_{kj}(t)\right]$$

$$= -\lambda(i) P_{ij}(t) + \lambda(i)\sum_{k\in S} Q_{ik} P_{kj}(t)$$

$$= \lambda(i)\left[-1\cdot P_{ij}(t) + \sum_{k\in S} Q_{ik} P_{kj}(t)\right]$$

$$= \lambda(i)\cdot\left[-(I\,P(t))_{ij} + (QP)_{ij}\right]$$

matrix multiplication



$$= \lambda(i)\left((-I + Q)P(t)\right)_{ij}$$

$$= \left(\Lambda(Q-I)P(t)\right)_{ij}$$

where $\Lambda = \operatorname{diag}(\lambda)$ is a diagonal matrix.

Thus, differentiating component by component,

$$\frac{dP(t)}{dt} = \Lambda(Q-I)P(t) = A\,P(t)$$

"Kolmogorov's backward equation"

where $A = \Lambda(Q-I)$ is called the *generator* of X.

Solution of the differential equation

$$\frac{dP(t)}{dt} = A\ P(t), \qquad P(0) = I$$

If there were real numbers instead of matrices, we would write

$$P(t) = e^{At}$$

where

$$e^{At} = \sum_{n=0}^{\infty} \frac{(At)^n}{n!} .$$

But the matrix A is square, and in fact everything we've written makes perfectly good sense here as well, if we make the (usual) convention that $(At)^0 = I$.

Then

$$P(0) = \frac{(A\cdot 0)^0}{0!} + \frac{(A\cdot 0)^1}{1!} + \cdots = I$$

$$\frac{d}{dt} P(t) = \frac{d}{dt} \sum_{n=0}^{\infty} \frac{(At)^n}{n!}$$

$$= \sum_{n=1}^{\infty} \frac{d}{dt} \frac{A^n \cdot t^n}{n!} \qquad t \text{ is a scalar}$$

$$= \sum_{n=1}^{\infty} \frac{A^n \cdot n t^{n-1}}{n!}$$

$$= \sum_{n=1}^{\infty} A \cdot \frac{A^{n-1} t^{n-1}}{(n-1)!} = \sum_{n=1}^{\infty} \frac{A^{n-1} t^{n-1}}{(n-1)!} \cdot A$$

$$= A\, e^{At} = e^{At} A ,$$

so that $P(t) = e^{At}$ <u>is</u> in fact a solution of $\frac{dP(t)}{dt} = A\,P(t)$. (in fact, the only solution).

Matlab

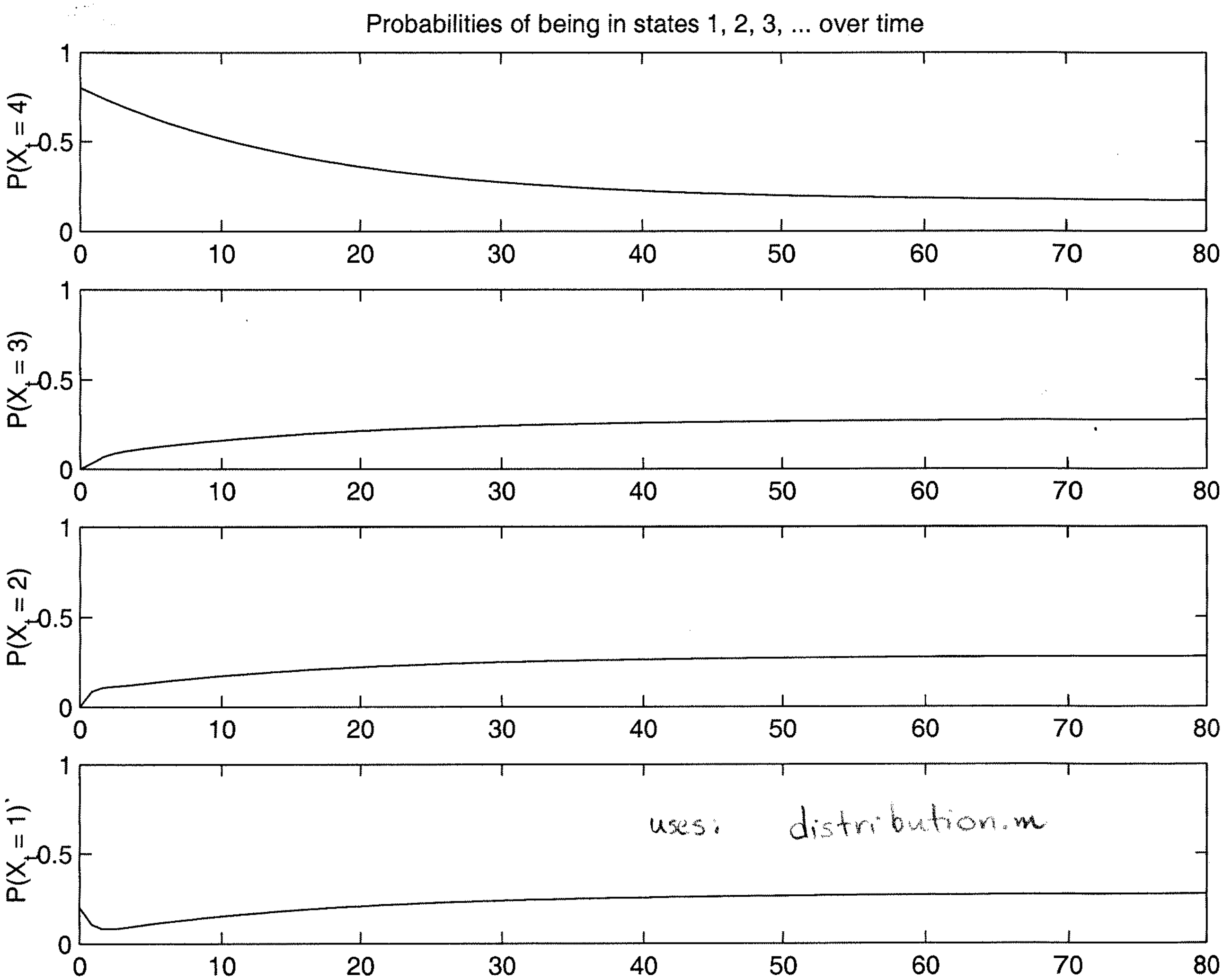
```
A = diag(lambda) * (Q - eye(size(Q)));
P(t) = expm(A*t);                 matrix exponential.
```

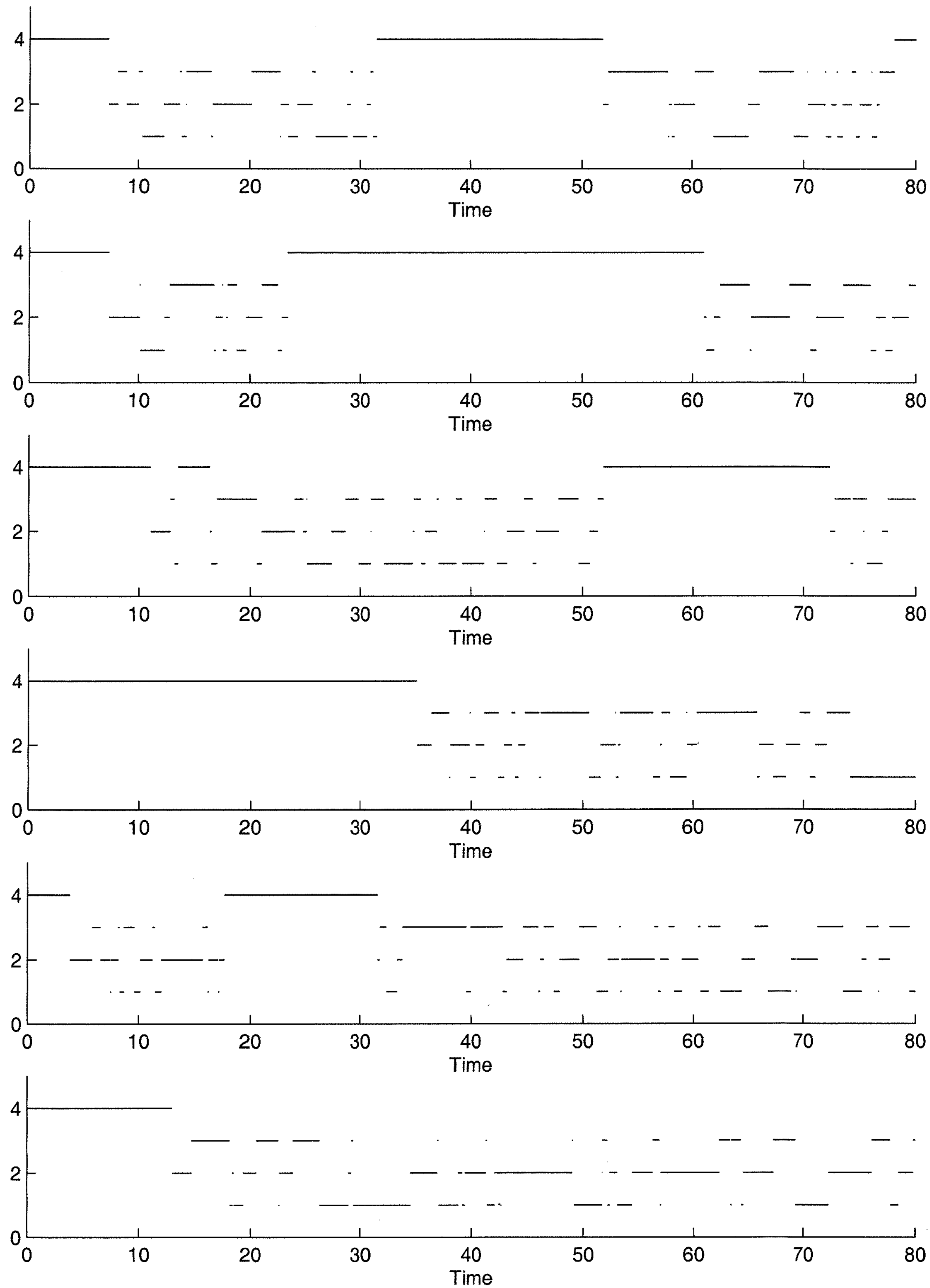
See the Markov process examples on my website.

```
using process_ex_2.m, we get:

mu =

    0.2000         0         0    0.8000

lambda =

    0.9050
    0.9050
    0.9050
    0.0500

Q =

         0    0.8840    0.1050    0.0110
    0.1050         0    0.8840    0.0110
    0.8840    0.1050         0    0.0110
         0    1.0000         0         0

>> A = diag(lambda)*(Q-eye(size(Q)))

A =

   -0.9050    0.8000    0.0950    0.0100
    0.0950   -0.9050    0.8000    0.0100
    0.8000    0.0950   -0.9050    0.0100
         0    0.0500         0   -0.0500

>> Pone = expm(A)

Pone =

    0.4713    0.3452    0.1738    0.0097
    0.1738    0.4715    0.3450    0.0097
    0.3450    0.1740    0.4713    0.0097
    0.0041    0.0331    0.0113    0.9515
```

Invariant distribution

$\eta = [0.2747 \quad 0.2812 \quad 0.2774 \quad 0.1667$





process_ex_2.m, then manyprocesses.m