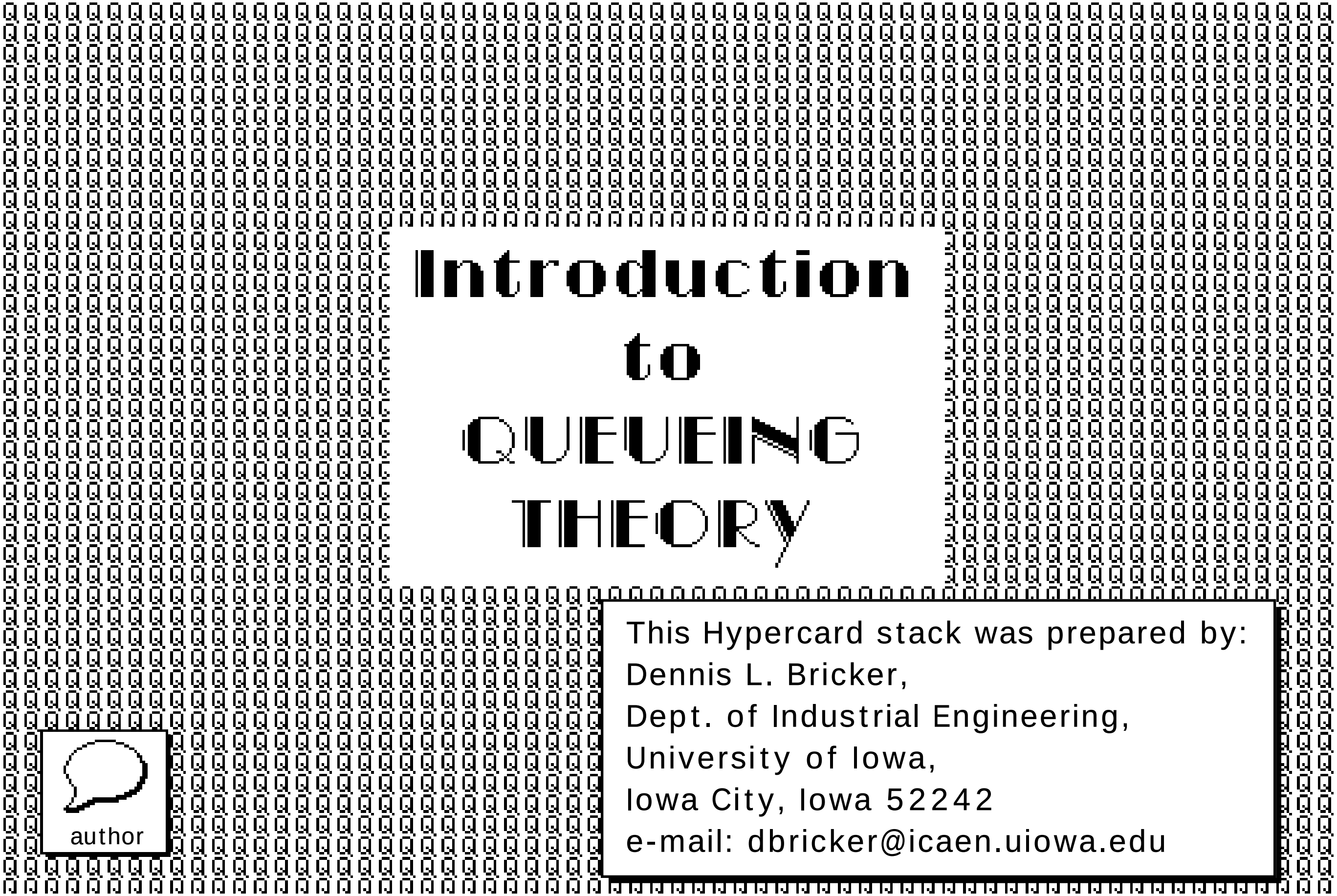


Introduction to QUEUEING THEORY

This Hypercard stack was prepared by:
Dennis L. Bricker,
Dept. of Industrial Engineering,
University of Iowa,
Iowa City, Iowa 52242
e-mail: dbricker@icaen.uiowa.edu

Kendall's Notation

Arrival Process / Service Process / Number of Servers / Capacity of System

(omitted if infinity)

M: Memoryless (Markovian)

E_k: Erlang-k

D: Deterministic

GI: General Interarrival times (but i.i.d.)

G: General service times (but i.i.d.)

The "Memoryless" arrival process indicates a Poisson arrival process, in which the interarrival times have an ***exponential*** distribution.

Likewise, the "Memoryless" service process indicates that the service times have an ***exponential*** distribution.

Some Markovian Queues

- ☞ M/M/1
- ☞ M/M/c (c>1)
- ☞ M/M/1/N
- ☞ M/M/1/N/N

Introduction
to
QUEUEING:
M/M/1

M/M/1

Interarrival times and service times both have exponential distributions, with parameters λ & μ, respectively.

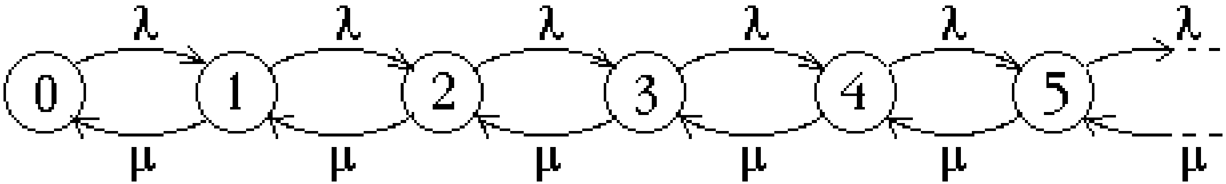
That is, the "customers" arrive at the rate of λ per unit time, and are served at the rate μ per unit of time.

It is assumed that the queue has infinite capacity, and that $\mu > \lambda$ (so that the queue length does not tend to increase indefinitely.)

In this case, it is possible to derive the probability distribution of the number of customers in the queueing system.

M/M/1

Birth/Death Model



$$\frac{1}{\pi_0} = 1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2 + \left(\frac{\lambda}{\mu}\right)^3 + \left(\frac{\lambda}{\mu}\right)^4 + \cdots$$

$$= \frac{1}{1 - \frac{\lambda}{\mu}}$$

(Geometric Series)

M/M/1

$\pi = (\pi_0, \pi_1, \pi_2, \ldots)$ denotes the "steady-state" distribution of the number of customers in this M/M/1 queueing system, i.e., 1+number in queue.

Equivalently, π_i is the probability (in steady state) that an arriving customer will find i customers already in the queueing system.

$$\boxed{\pi_i = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^i}$$

PASTA

Poisson
Arrivals
See
Time
Averages

PASTA implies that if we observe the system at any random point in time, the probability of the state that we observe will equal the fraction of the time that the system spends in that state!

M/M/1

$$\boxed{\pi_i = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^i}$$

$$\pi_0 = 1 - \frac{\lambda}{\mu}$$

Probability that no customers are in the system (server is idle)

Probability that the server is busy

$$1 - \pi_0 = \frac{\lambda}{\mu} \equiv \rho < 1$$

utilization

M/M/1

Average number of "customers" in the system (including any being served)

$$L = \sum_{j=0}^{\infty} j\,\pi_j = \sum_{j=0}^{\infty} j\,\underbrace{\left(1-\rho\right)\rho^j}_{\pi_j} = \left(1-\rho\right)\rho\underbrace{\sum_{j=0}^{\infty} j\,\rho^{j-1}}$$

Next we will derive an expression for this infinite sum!

where $\rho = \frac{\lambda}{\mu} < 1$

Derivative of a Geometric Series

This is not a geometric series, because of the appearance of the factor "j", but it is the sum of the derivatives of the terms of a geometric series!

$$\sum_{j=0}^{\infty} j\,\rho^{j-1} = \frac{d}{d\rho} \underbrace{\sum_{j=0}^{\infty} \rho^{j}} = \frac{d}{d\rho}\left(\frac{1}{1-\rho}\right) = \frac{1}{(1-\rho)^2}$$

Geometric series, with sum $\left(\frac{1}{1-\rho}\right)$

sum of derivatives equals derivative of the sum!

M/M/1

Average number of "customers" in the system (including any being served)

$$\boxed{L = \frac{\frac{\lambda}{\mu}}{1 - \frac{\lambda}{\mu}}} = \frac{\rho}{1 - \rho}$$

This formula ONLY for M/M/1 queue!

LITTLE's Queueing Formula

$$L = \underline{\lambda}\ W$$

average number in the queueing system

average arrival rate

average time in system per customer

☞ applies to ***any*** queueing system having a steady state distribution ***(whether Markovian or not!)***

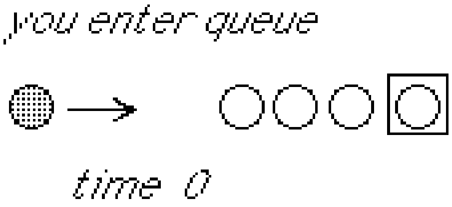
LITTLE's Queueing Formula

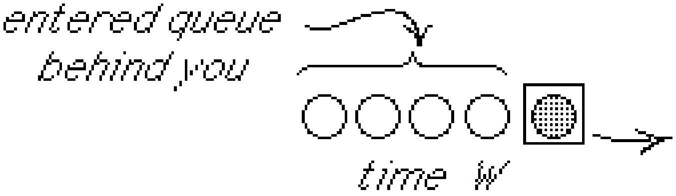
$$L = \underline{\lambda}\, W$$

Intuitive argument:

Suppose that you join a queue and spend **W** minutes before you have been served and leave.

During those **W** minutes, customers have been arriving and joining the queue behind you at the average rate of $\underline{\lambda}$ per minute. Thus, when you are ready to leave, you should expect to see $\underline{\lambda}$**W** customers remaining in the system behind you.





M/M/1

For the M/M/1 queueing system, Little's formula implies that

$$W = \frac{L}{\lambda} = \frac{\rho}{\lambda(1-\rho)} \qquad \textit{where} \quad \rho = \lambda/\mu$$

$$\Longrightarrow \boxed{W = \frac{1}{\mu - \lambda}}$$

Denote

L_q = average number of customers in the queue (not including those being served)

W_q = average time spent per customer in the queue (not including service time)

Then for any queueing system having a steady state distribution,

$$W_q = W - \frac{1}{\mu}$$

average rate of service

$$L_q = \lambda\, W_q$$

Little's Formula

Thus, given any one of the quantities

$$L, \; W, \; L_q, \text{ and } W_q$$

we can compute the others.

(Generally, L is the easiest to compute first!)

M/M/1

For the M/M/1 queueing system, then

$$W_q = W - \frac{1}{\mu} \Longrightarrow \boxed{W_q = \frac{\lambda}{\mu(\mu - \lambda)}}$$

$$L_q = \lambda\, W_q \Longrightarrow \boxed{L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}}$$

Example

An average of 24 trucks per 8-hour day arrive to be unloaded &/or loaded, which requires an average of 15 minutes.

The loading dock can handle only a single truck at a time.

Assume that the arrival process is Poisson, and that the service times have exponential distribution.

This loading dock is modeled as an M/M/1 queue.

M/M/1

λ = arrival rate = 3/hour

μ = service rate = 4/hour

Utilization of the server

$$\rho = \frac{\lambda}{\mu} = 0.75$$

Average number of trucks in system

$$L = \frac{\rho}{1 - \rho} = \frac{0.75}{1-0.75} = 3$$

Average time in system per truck

$$W = \frac{L}{\lambda} = \frac{3}{3/hr} = 1 \text{ hr.}$$

Steady-state Behavior

M/M/1

λ = arrival rate = 3/hour

μ = service rate = 4/hour

Average time in the queue

$$W_q = W - \frac{1}{\mu} = 1 \text{ hr.} - \frac{1}{4/\text{hr}}$$

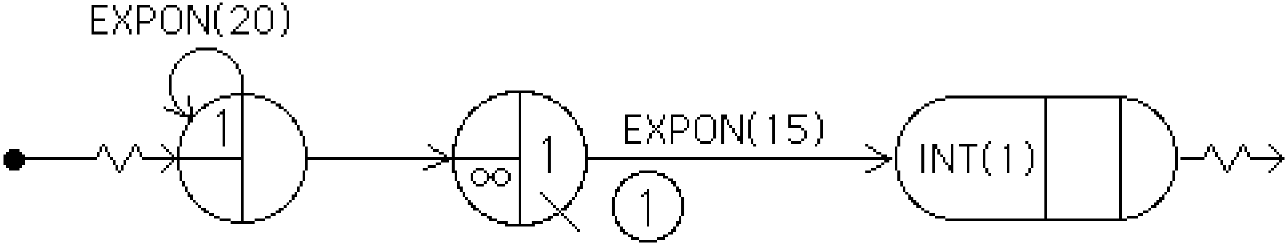
$$= 0.75 \text{ hr.}$$

Average length of the queue

$$L_q = \underline{\lambda}\, W_q = (3/\text{hr})(0.75\text{hr})$$

$$= 2.25$$

SLAM Model



M/M/1 Queue

SLAM code

```
GEN,BRICKER,MM1_QUEUE,3/18/92,1,Y,Y,Y/N,Y,Y,72;
LIM,1,2,100;
INIT,,14400;                    SIMULATE TEN 24-HOUR DAYS
NETWORK;
        CREATE,EXPON(20),,1;    ARRIVAL OF TRUCKS
        QUEUE(1);
        ACTIVITY(1)/1,EXPON(15);
        COLCT,INT(1),TIME IN SYSTEM,20/10/10;
        TERM;
        END;
FIN;
```

```
              **STATISTICS FOR VARIABLES BASED ON OBSERVATION**

               MEAN      STANDARD  COEFF. OF  MINIMUM   MAXIMUM   NO.OF
               VALUE     DEVIATION VARIATION   VALUE    VALUE     OBS
     TIME IN
     SYSTEM  0.629E+02 0.650E+02 0.103E+01 0.430E-01 0.327E+03  741
```

W

```
                **FILE STATISTICS**


FILE     LABEL/   AVERAGE  STANDARD   MAXIMUM  CURRENT   AVERAGE
NUMBER    TYPE    LENGTH   DEVIATION  LENGTH   LENGTH    WAIT TIME
   1     QUEUE      2.499     3.640      20        1     48.431
   2     CALENDAR   1.742     0.438       3        2     12.290
```

L_q

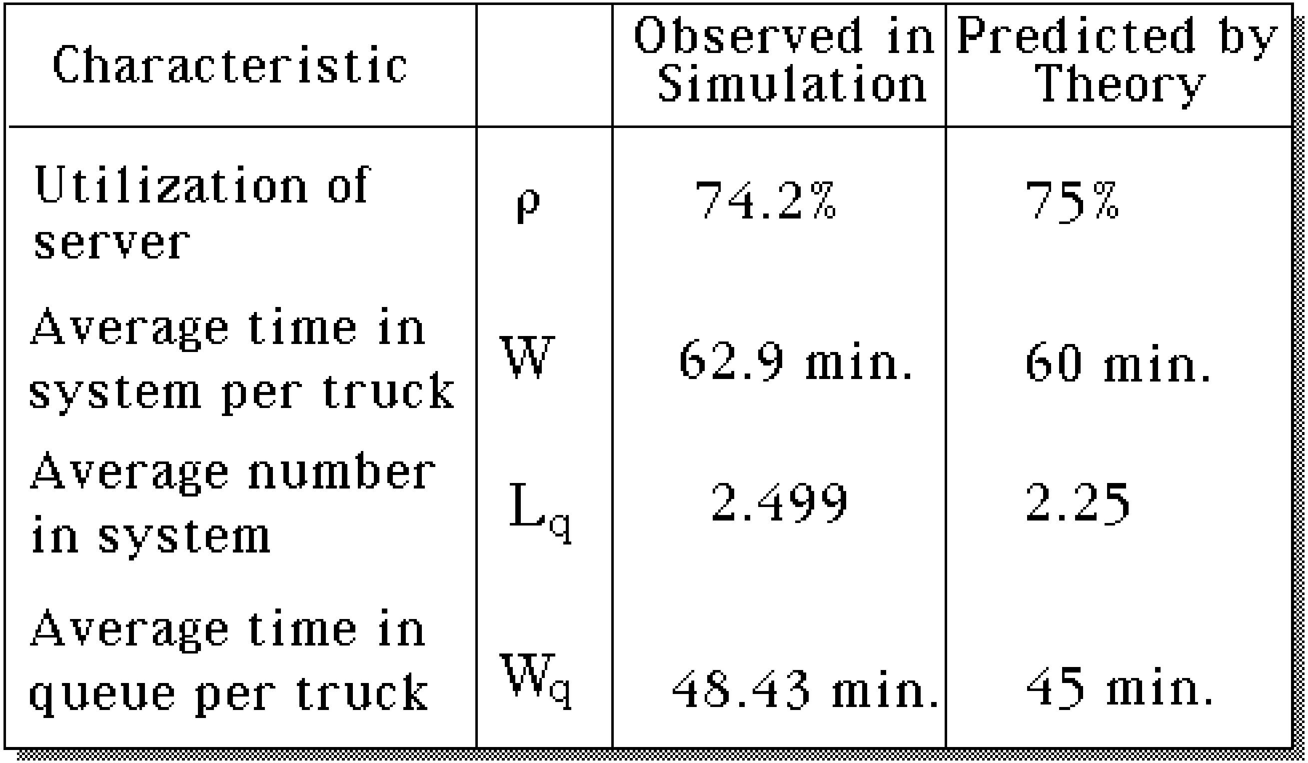
```
            **SERVICE ACTIVITY STATISTICS**

ACT  SER AVERAGE   STD  CUR  AVERAGE  MAX IDL   MAX BSY  ENT
NUM  CAP   UTIL    DEV  UTIL  BLOCK   TME/SER   TME/SER  CNT

  1    1  0.742   0.44    1    0.00    107.23   1557.37  741
```

ρ

```
OBS  RELA   UPPER
FREQ FREQ  CELL LIM 0        20        40        60        80       100
                    +    +    +    +    +    +    +    +    +    +    +
122 0.165 0.100E+02 +********                                         +
103 0.139 0.200E+02 +*******       C                                  +
 70 0.094 0.300E+02 +*****              C                             +
 62 0.084 0.400E+02 +****                   C                         +
 63 0.085 0.500E+02 +****                       C                     +
 55 0.074 0.600E+02 +****                           C                 +
 43 0.058 0.700E+02 +***                               C              +
 24 0.032 0.800E+02 +**                                  C            +
 28 0.038 0.900E+02 +**                                   C           +
 17 0.023 0.100E+03 +*                                      C         +
 20 0.027 0.110E+03 +*                                       C        +
 18 0.024 0.120E+03 +*                                        C       +
  7 0.009 0.130E+03 +                                          C      +
 16 0.022 0.140E+03 +*                                          C     +
 10 0.013 0.150E+03 +*                                          C     +
 11 0.015 0.160E+03 +*                                           C    +
  3 0.004 0.170E+03 +                                            C    +
  9 0.012 0.180E+03 +*                                            C   +
  7 0.009 0.190E+03 +                                             C   +
  9 0.012 0.200E+03 +*                                             C  +
  9 0.012 0.210E+03 +*                                              C +
 35 0.047    INF    +**                                               C
---                 +    +    +    +    +    +    +    +    +    +    +
741                 0        20        40        60        80       100
```



Characteristic		Observed in Simulation	Predicted by Theory
Utilization of server	ρ	74.2%	75%
Average time in system per truck	W	62.9 min.	60 min.
Average number in system	L_q	2.499	2.25
Average time in queue per truck	W_q	48.43 min.	45 min.

, U. of Iowa, 1997

Example The arrival process at a certain work center is Poisson, with rate 3/hour.

The service time has exponential distribution, with mean 0.3 hr.

Each job waiting for processing requires 1 m^2 of floor space.

How much in-process storage space should be allocated to accomodate all waiting jobs

... 90% of the time?

... 95% of the time?

... 99% of the time?

If **n** square meters of floor space are allocated, this will be sufficient a fraction of time equal to

$$P\{\text{number in system} \le n+1\} = \sum_{j=0}^{n+1} \pi_j$$

$$\boxed{\pi_j = \rho^j (1 - \rho)}$$

$$\Longrightarrow \quad \sum_{j=0}^{n+1} \pi_j = \sum_{j=0}^{n+1} \rho^j (1 - \rho)$$

sum of finite geometric series

$$= (1 - \rho) \sum_{j=0}^{n+1} \rho^j = (1 - \rho) \left[\frac{1 - \rho^{n+2}}{1 - \rho} \right]$$

$$= 1 - \rho^{n+2}$$

Solving for n: $1 - \rho^{n+2} \geq \alpha$

$$\Longrightarrow \quad \rho^{n+2} \leq 1 - \alpha$$

$$(n+2) \log \rho \leq \log (1 - \alpha)$$

$$n+2 \geq \frac{\log (1 - \alpha)}{\log \rho}$$

$$\boxed{n \geq \frac{\log (1 - \alpha)}{\log \rho} - 2}$$

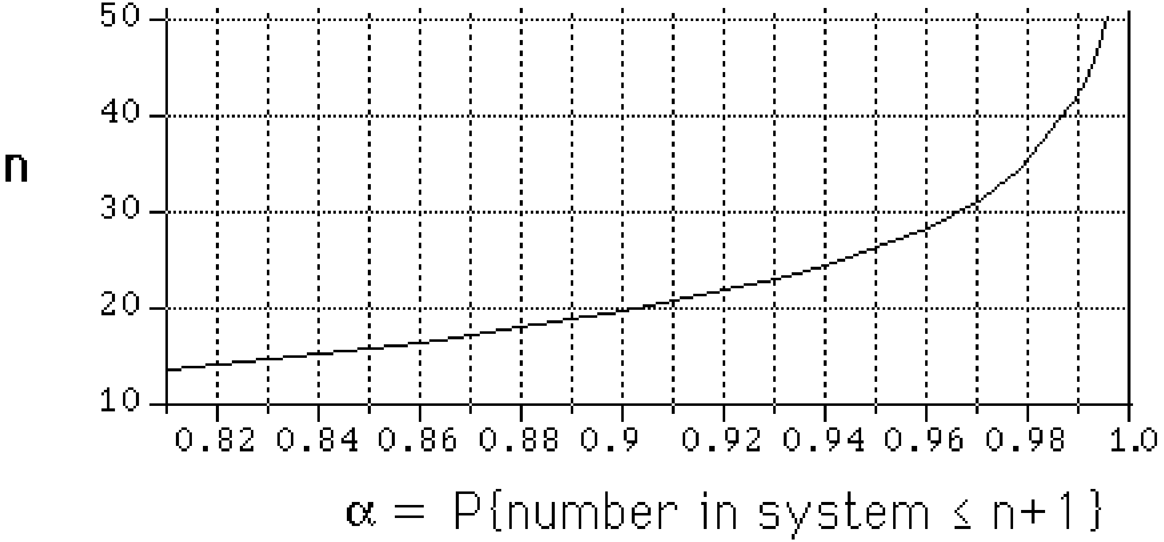
& integer!

$$n \geq \frac{\log(1-\alpha)}{\log \rho} - 2$$

where

$\rho = 0.9$

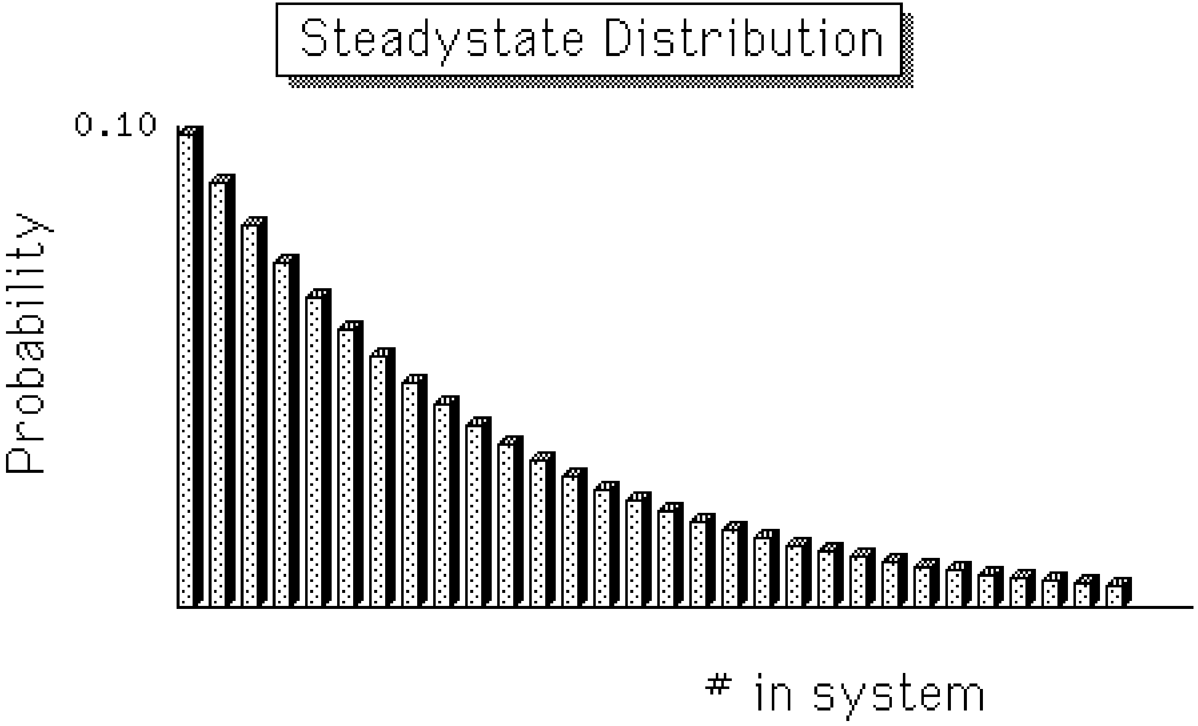
α	$\frac{\log(1-\alpha)}{\log \rho} - 2$	n
0.85	16.00598614	17
0.9	19.85434533	20
0.95	26.43315881	27
0.96	28.5510637	29
0.97	31.28151799	32
0.98	35.12987717	36
0.99	41.70869065	42
0.992	43.82659554	44
0.994	46.55704984	47
0.996	50.40540902	51


n
50
40
30
20
10
0.82
0.84
0.86
0.88
0.9
0.92
0.94
0.96
0.98
1.0
α = P{number in system ≤ n+1}

Steadystate Distribution

i	Pi	CDF
0	0.100000	0.100000
1	0.090000	0.190000
2	0.081000	0.271000
3	0.072900	0.343900
4	0.065610	0.409510
5	0.059049	0.468559
6	0.053144	0.521703
7	0.047830	0.569533
8	0.043047	0.612580
9	0.038742	0.651322
10	0.034868	0.686189
11	0.031381	0.717570
12	0.028243	0.745813
13	0.025419	0.771232
14	0.022877	0.794109
15	0.020589	0.814698

i	Pi	CDF
15	0.020589	0.814698
16	0.018530	0.833228
17	0.016677	0.849905
18	0.015009	0.864915
19	0.013509	0.878423
20	0.012158	0.890581
21	0.010942	0.901523
22	0.009848	0.911371
23	0.008863	0.920234
24	0.007977	0.928210
25	0.007179	0.935389
26	0.006461	0.941850
27	0.005815	0.947665
28	0.005233	0.952899
29	0.004710	0.957609
30	0.004239	0.961848


Steadystate Distribution
0.10
Probability
in system

Cumulative Steady-State Probabilities