G. GIALLOMBARDO

3E3 Modelling Risk
Exam Paper 2003

Crib

1(a)

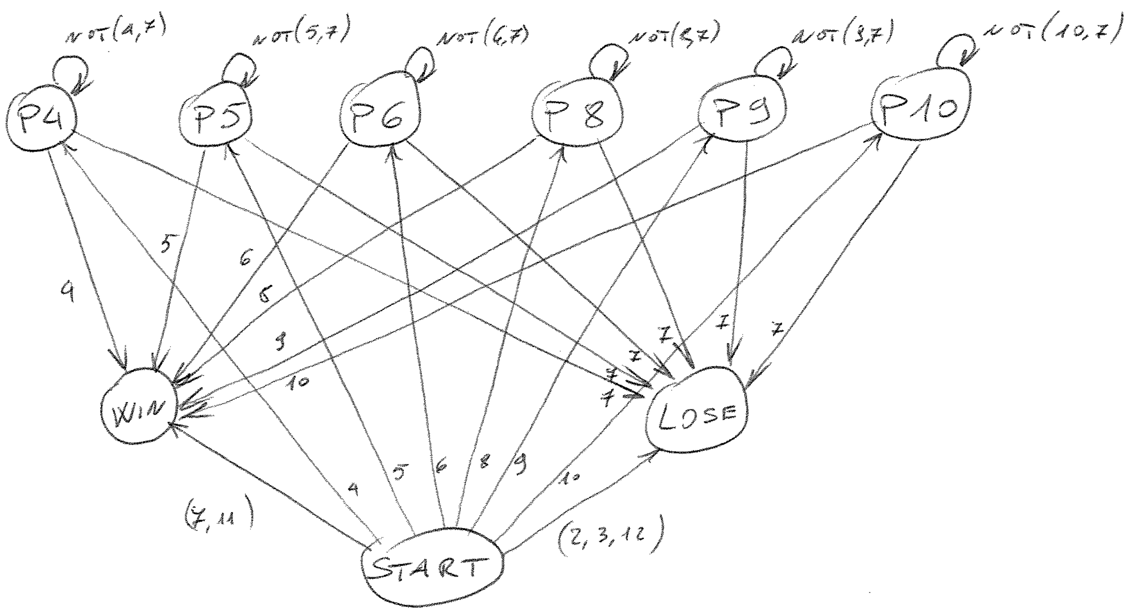
The state transition network is shown below. The events are the dice rolls (arcs in the graph). The states (nodes in the graph) represent the possible outcomes of each dice roll (win, lose, or a 'point'). The player begins at the node 'Start'. The first roll leads to 'Win', 'Lose', or one of the states representing a 'point' (P4, P5, P6, P8, P9, P10). The process stops when the system enters state 'Win' or 'Lose'.

The transition probability are calculated by simple probability analysis. For instance the transition probability from state 'Start' to state 'Win' is the probability of getting 7 or 11. Since there are six possible ways of getting 7, and two possible ways of getting 11, out of 36 total possible outcomes, it is easy to calculate the transition probability as 8/36=0.222.

Transition Matrix:

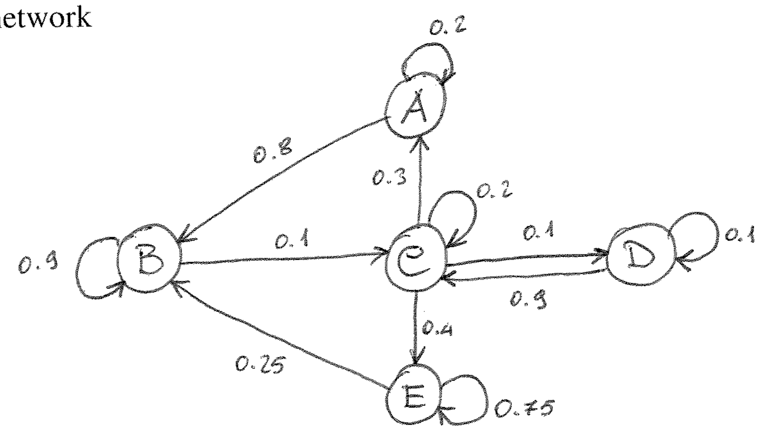
From / To	Start	Win	Lose	P4	P5	P6	P8	P9	P10
Start	0	0.222	0.111	0.083	0.111	0.139	0.139	0.111	0.083
Win	0	1	0	0	0	0	0	0	0
Lose	0	0	1	0	0	0	0	0	0
P4	0	0.083	0.167	0.75	0	0	0	0	0
P5	0	0.111	0.167	0	0.722	0	0	0	0
P6	0	0.139	0.167	0	0	0.694	0	0	0
P8	0	0.139	0.167	0	0	0	0.694	0	0
P9	0	0.111	0.167	0	0	0	0	0.722	0
P10	0	0.083	0.167	0	0	0	0	0	0.75

Transition Network



1(b)(i)
There is only one class (absorbing of course) containing all the states.

Transition network



1(b) (ii)
Given q(0)=[0,1,0,0,0], it is possible to calculate q(1)=q(0)P and then
q(2)=q(1)P=[0.03, 0.81, 0.11, 0.01, 0.04]
and
q(3)=q(2)P=[0.039, 0.763, 0.102, 0.020, 0.074].

1(b) (iii)
Yes, because the Markov chain has finite state space and it is irreducible and aperiodic.
To calculate the limiting distribution vector it is sufficient to solve the following linear system
$\pi=\pi P$
$\pi_1+\pi_2+\pi_3+\pi_4+\pi_5=1$

from which we obtain
$\pi_1= 0.034$
$\pi_2= 0.638$
$\pi_3= 0.091$
$\pi_4= 0.091$
$\pi_5= 0.145$

2(a)

U = Interarrival distribution;
V = Service time distribution;
s = Number of servers;
k = Maximum number of customer in the system (in service and in queue);
W = Queuing discipline.

Assumptions:

- Interarrival times are independent and identically distributed.
- Service times are independent and identically distributed.
- Interarrival and service times are independent.

2(b)

The state N(t) of a queuing system at time t is the number of customers in the system (i.e. in the queue or in service) at time t.
The system is said to be in steady state if P(N(t)=n) does not change with t any more.
If we call λ and μ respectively the arrival rate and the service rate in a G/G/s queue, the utilization factor is the ratio $\rho=\lambda/(s\mu)$.
ρ is the fraction of time we expect the service facility to be busy. Therefore if $\rho>1$ the queue explodes. In fact the condition $\rho<1$ is necessary for the existence of a steady state solution.

2(c)

The random variable X has the lack of memory property if for any t,s>0 it holds:
$P(X>t+s \mid X>s)=P(X>t)$.
The lack-of-memory property is typical for interarrival times, provided arrival rates are constant over time. But, if a continuous random variable (with a continuous density function) has the lack-of-memory property then it is exponentially distributed. This is the reason for the prevalence of the exponential distribution in queuing models.

2(d)

By applying the birth and death process theory:
For the M/M/1 queue with arrival rate λ and service rate μ, we have $\lambda_n=\lambda$, $\mu_n=\mu$ for all n.

Hence $c_n=(\lambda/\mu)^n=\rho^n$ and $s=c_1+c_2+...=\rho^1+\rho^2+...=\rho/(1-\rho)$.

Hence $p_0=1/(1+s)=1-\rho$, $p_n=c_n p_0=(1-\rho)\rho^n$.

Thus $L=1p_1+2p_2+3p_3+...=(1-\rho)(1\rho+2\rho^2+3\rho^3+...)=\rho/(1-\rho)=\lambda/(\mu-\lambda)$,
and $L_q=1p_2+2p_3+3p_4+...=\rho^2/(1-\rho)$.
Moreover using Little's formula and the relation $W=W_q+1/\mu$, we have
$W=1/\mu(1-\rho)$ and $W_q=\rho/\mu(1-\rho)$.

2(e)

Let A(t) denote the number of customers that arrive in the time interval [0,t]. If the interarrival times are independent and exponentially distributed with parameter λthen A(t) has a Poisson distribution with parameter λt, i.e.

$$P(A(t)=n)=e^{-\lambda t}((\lambda t)^n/n!$$

and A(t) is called a Poisson arrival process.
$E[A(t)]=V[A(t)]=\lambda t$ (mean=variance).

3(a)

The sum of squared errors is of the form

$$SSE(a,b)=\sum_{i=1}^{n}(a+bx_i-y_i)^2 .$$

Differentiating this with respect to the two the two parameters gives

$$\frac{\partial SSE}{\partial a}(a,b)=\sum 2(a+bx_i-y_i)=2na+2b\sum x_i-2\sum y_i=0$$

$$\frac{\partial SSE}{\partial b}(a,b)=\sum 2(a+bx_i-y_i)x_i=2a\sum x_i+2b\sum x_i^2-2\sum y_i x_i=0$$

This is a linear system in the two unknowns a,b. Dividing the first equation by $2n$ gives

$$a+b\bar{x}=\bar{y},$$

where $\bar{x}=\sum x_i/n$, $\bar{y}=\sum y_i/n$ are the averages of the x and y values in the data. Thus the line of best fit passes through the "average point" $(\bar{x},\bar{y})$. Plugging this into the second equation gives

$$b=\frac{\sum x_i(y_i-\bar{y})}{\sum x_i(x_i-\bar{x})}=\frac{\sum(x_i-\bar{x})y_i}{\sum(x_i-\bar{x})^2}=\sum\lambda_i y_i,$$

where $\lambda_i=(x_i-\bar{x})/\sum(x_j-\bar{x})^2$,

and

$$a=-b\bar{x}+\bar{y}.$$

3(b)

Since the observations y_i are contaminated with errors, the least squares estimates depend on the sample produced by the random data generating process. Therefore the line of best fit is a random line.

3(c)

Recall that $b=\sum\lambda_i(\alpha+\beta x_i+\varepsilon_i)=\underbrace{\sum\lambda_i(\alpha+\beta x_i)}_{Non-random}+\underbrace{\sum\lambda_i\varepsilon_i}_{Random}.$

The random element is the sum of n random variables $\lambda_i\varepsilon_i$, where n is the number of data points (x_i,y_i). We know from the central limit theorem that the sum of independent random variables (whatever their distribution) tends to become more and more normal as the number of random variables increases. If the number n is large enough, say $n\geq 30$ as a rule of thumb, then we may safely operate under the assumption that the sum is indeed normal. The expectation is obviously zero since the expected value of a linear combination of random variables is the linear combination of the expected values. For the calculation of the variance we need to use the fact that the variance of the sum of independent variables is the sum of their variances and that $VAR(aX)=a^2VAR(X)$ for any number a and any random variable X. By applying these two rules we obtain

$$VAR(\sum\lambda_i\varepsilon_i)=\sigma^2\sum\lambda_i^2=\frac{\sigma^2}{\sum(x_i-\bar{x})^2},$$

where σ^2 is the common variance of the error terms ε_i. The last equation is a direct consequence of the definition of the coefficients λ_i.

Let us now investigate the non-random term $\sum \lambda_i(\alpha + \beta x_i)$ in our formula for the least squares slope b, which is of course the expected value of b, since the random term has expected value zero. We have

$$\sum \lambda_i(\alpha + \beta x_i) = \alpha(\sum \lambda_i) + \beta(\sum \lambda_i x_i).$$

Notice first that

$$\sum \lambda_i = \frac{\sum (x_i - \bar{x})}{\sum (x_j - \bar{x})^2} = \frac{(\sum x_i) - n\bar{x}}{\sum (x_j - \bar{x})^2} = \frac{(\sum x_i) - (\sum x_i)}{\sum (x_j - \bar{x})^2} = 0.$$

Furthermore,

$$\sum \lambda_i x_i = \sum \lambda_i x_i - \bar{x}\sum \lambda_i = \sum \lambda_i (x_i - \bar{x}) = \frac{\sum (x_i - \bar{x})(x_i - \bar{x})}{\sum (x_j - \bar{x})^2} = 1$$

and therefore

$$\sum \lambda_i(\alpha + \beta x_i) = \beta.$$

Consequently: Our analysis shows that the least squares slope b has (approximately) a normal distribution with mean β and variance $\sigma^2 / \sum (x_i - \bar{x})^2$.

3(d)

$a = 95.612$
$b = 1.047$

4(a)
P(all servers idle)= $\pi_0 = 0.068$

4(b)
P(no wait)= $\pi_0 + \pi_1 + \pi_2 = 0.450$

4(c)
P(wait)=1 - P(no wait) – P(full) = $1 - (\pi_0 + \pi_1 + \pi_2 + \pi_6) = 0.447$

4(d)
P(lost customer) = $\pi_6 = 0.103$

4(e)
$L_q = 1\pi_4 + 2\pi_5 + 3\pi_6 = 0.702$

4(f)
$L_s = \pi_1 + 2\pi_2 + 3(1 - \pi_0 - \pi_1 - \pi_2) = 2.244$

4(g)
$E = L_s / s = 2.244 / 3 = 74.8\%$