Solutions to some Continuous Stochastic Modelling exam questions

May 2002, question 2

(i) Assume a Poisson process. Rate is 2 per hour. The probability of no arrivals in a 1-hour period is e^{-2}.

(ii) Using the independent increments property, we see that the probability of getting more than 2 messages between 2.00 and 5.00 is independent of how many messages arrived during the lunch hour. The required probability is $1 - e^{-6}(1 + 6 + 6^2/2)$, since the number of messages arriving in a 3-hour period has a Poisson distribution with mean 6.

(iii) The number of messages arriving between 2.00 and 5.00 has Poisson(6) distribution. Each one, independently, has probability $\frac{1}{3}$ of being the kind which cannot be dealt with by Admissions Secretaries. Therefore this is a thinned Poisson process: the number which can't be dealt with by the ASs has Poisson(2) distribution, and the probability that the number is less than 3 is $e^{-2}(1 + 2 + 2^2/2)$.

May 2002, question 6

(i) Bookwork.

(ii) a) The Markov property ensures that the probability of remaining in state 3 continuously from time 2 to time 4 is independent of anything that happened before time 2, given that $X(2) = 3$. The required probability is $\exp[-(4-2)\lambda_3] = e^{-0.8}$.

b) Here we require

$$\int_2^4 \lambda_3 e^{-\lambda_3(t-2)} \times r_{34}\, dt = (1 - e^{-0.8}) \times \frac{1}{4}.$$

c) We are looking for an expression for $p_{13}(t)$. By conditioning on the time of the first transition out of state 1, we can write this as

$$\begin{aligned} p_{13}(t) &= \int_0^t \lambda_1 e^{-\lambda_1 u}(r_{12}p_{23}(t-u) + r_{14}p_{43}(t-u))\, du \\ &= \int_0^t e^{-0.5u}(0.4p_{23}(t-u) + 0.1p_{43}(t-u))\, du \end{aligned}$$

d) We condition on the first transition taking place at time u, or, more strictly, between time u and $u + du$. The probability of this is $f_T(u)\, du$, where T is the random variable representing the time of the first transition out of state 1. We know that T has exponential distribution, rate 0.5, so this element of the equation is equal to $0.5e^{-0.5u}\, du$.

This must be multiplied by the probability of being in state 3 at time t given that the first transition out of state 1 takes place at time u. At this point we must condition again, but this time on the destination of the first jump out of state 1. The first jump takes the process to state 2 with probability $r_{12} = 0.8$, or to state 4 with probability $r_{14} = 0.2$.

The final factor is the conditional probability, given that the first transition takes place at time u and is to state k, that the process is in state 3 at time t. This is equal to $p_{k3}(t-u)$.

June 2003 Question 1 The estimate of q_{EF} is

$$\hat{q}_{EF} = \frac{\text{Number of transitions from } E \text{ to } F}{\text{Total time spent in } E},$$

with $\hat{q}_{FE}$ determined similarly.

June 2003 Question 4

(i) It is reasonable to suppose that the times between calls are independent random variables, but perhaps less reasonable to imagine that they occur at the same rate throughout the day. It is possible that a time-inhomogeneous Poisson process may be a better model.

(ii) The arrivals of calls that cannot be answered in-house forms a thinned Poisson process, so is itself a Poisson process with rate 12 per hour. Thus the number of calls arriving in a 5-minute period has Poisson(1) distribution, so the required probability is $e^{-1}(1+1+1/2)$.

(iii) By the independent increments property, the number of calls arriving between 1.30 and 2.30 is irrelevant to the question; we require only the probability that a Poisson(2) random variable takes the value 1, which is $2e^{-2}$.

June 2004 Question 6

(i) $1/\lambda_D = 1/\rho$.

(ii) $e^{-\rho t}$.

(iii) We require

$$\int_3^4 f_T(u) r_{DA}\, du,$$

where $f_T(u)$ is the density function of the time of first leaving state D, i.e. $f_T(u) = \rho e^{-\rho u}$, and r_{DA} is the conditional probability that the first transition from D takes the process to state A, which is 1 (since there are no other states). Therefore we need

$$\int_3^4 \rho e^{-\rho u}\, du = e^{-3\rho} - e^{-4\rho}.$$

(iv) The KFE is

$$\frac{d}{dt} p_{DD}(t) = -\rho p_{DD}(t) + \mu p_{DA}(t).$$

(v) We note that $p_{DA}(t) = 1 - p_{DD}(t)$. Therefore

$$\frac{d}{dt} p_{DD}(t) + (\rho+\mu) p_{DD}(t) = \mu.$$

The integrating factor is $\exp((\rho+\mu)t)$, so that we have

$$\frac{d}{dt}\left(e^{(\rho+\mu)t} p_{DD}(t)\right) = \mu e^{(\rho+\mu)t},$$

and integrating both sides gives us

$$e^{(\rho+\mu)t} p_{DD}(t) - p_{DD}(0) = \frac{\mu}{\rho+\mu}\left(e^{(\rho+\mu)t} - 1\right).$$

Applying the boundary condition $p_{DD}(0) = 1$ gives the answer.

(vi) The information supplied implies that $\rho = 0.2$ per hour, $\mu = 1$ per hour. We require

$$p_{DA}(2) = 1 - p_{DD}(2) = \frac{\rho}{\mu+\rho}\left(1 - e^{-2(\mu+\rho)}\right) = \frac{1}{6}\left(1 - e^{-2.4}\right).$$

(vii) We need an expression for $p_{DD}(s,t)$ in integral form. Using standard arguments, the integrated form of the KFE is

$$p_{DD}(s,t) = \exp\left(-\int_s^t \rho(u)\,du\right) + \int_s^t \rho(u)\exp\left(-\int_s^u \rho(v)\,dv\right) p_{AD}(u,t)\,du,$$

but it is possible that the question intends us to condition on the time of the *last* transition from A to D, in which case we should have

$$p_{DD}(s,t) = \exp\left(-\int_s^t \rho(u)\,du\right) + \int_s^t p_{DA}(s,u)\mu(u)\exp\left(-\int_u^t \rho(v)\,dv\right) du.$$

January 2004 Question 1

(i) The instantaneous mortality rate, $\mu(t)$, is defined by

$$\mu(t) = \lim_{dt\to 0} \frac{\mathbb{P}[X(t+dt) = D \mid X(t) = A]}{dt}$$

(ii) The required solution is

$$p_{AD}(4,6) = 1 - \exp\left(-\int_4^6 \mu(t)\,dt\right) = 1 - \exp\left(-\int_4^6 0.03e^{0.5t}\,dt\right) = 1 - \exp(\left(-0.06(e^3 - e^2)\right).$$

January 2004 Question 3

(i) The *residual holding time* in state i at time t is defined as $T - t$, where T is the time of the next transition out of state i.

(ii) $\exp\left(-\int_s^t (\sigma(u) + \mu(u))\,dt\right)$.

(iii) Assume $X(s) = H$ and define T as the time which passes from s until the first transition out of state H. Then the probability calculated in (ii) is $\mathbb{P}(T > t)$, which implies that the distribution function of T is

$$F_T(t) = 1 - \exp\left(-\int_s^t (\sigma(u) + \mu(u))\,dt\right),$$

and simple differentiation implies that

$$f_T(t) = \frac{d}{dt}F_T(t) = (\sigma(t) + \mu(t))\exp\left(-\int_s^t (\sigma(u) + \mu(u))\,dt\right)$$

for $t > s$.

January 2004 Question 6

(i) There are a number of elements that could be tested, such as whether the distribution of time spent in each state is exponential, or whether the state to which the process jumps after a transition from state i is independent of the time spent in state i before the transition occurred.

Each of these has a specific test, but if you want to test all aspects at once, a good procedure is to use the data to estimate transition rates and to simulate the process a large number of times. If the simulated data looks pretty much the same as the real data, then the model is a good one.

(ii) The generator matrix can be estimated as

$$\hat{Q} = \begin{pmatrix} -\frac{1}{2} & \frac{1}{8} & 0 & 0 & \frac{3}{8} \\ 0 & -\frac{1}{7} & \frac{9}{70} & 0 & \frac{1}{70} \\ 0 & 0 & -\frac{1}{14} & \frac{1}{140} & \frac{9}{140} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

(iii) Since the distribution is exponential with rate $\frac{1}{7}$ per day, the probability is $\exp(-10/7)$.

(iv) Since the process is time-homogeneous, it doesn't matter how long ago the references were first sought. We need

$$\int_8^{10} \frac{1}{7} e^{-t/7} r_{BC}\, dt = \frac{9}{10}(e^{-8/7} - e^{-10/7}).$$

(v) The first phase is waiting for references, which takes an average of 7 days. With probability 10% that is the end of the matter, but with probability 90% the candidate has to wait for interview, which takes an average of 14 days. Thus the total expected time is $7 + 0.9 \times 14 = 19.6$ days.

(vi) Condition on the time, u, at which the reference is received. We need to multiply together the density function $f_T(u)$, the probability that the reference is favourable, and the probability that the interview has been sorted out by time 17 given that references are received at time u. This is

$$\int_0^{17} \frac{1}{7} e^{-u/7} \times \frac{9}{10} \times e^{-(17-u)/14}\, du = \frac{9}{70} e^{-17/14} \int_0^{17} e^{-u/14}\, du = \frac{9}{5}\left(e^{-17/14} - e^{-17/7}\right).$$

(vii) The KBEs are

$$\begin{aligned} \frac{dp_{AC}}{dt} &= q_{AA}p_{AC}(t) + q_{AB}p_{BC}(t) = -\frac{1}{2}p_{AC}(t) + \frac{1}{8}p_{BC}(t) \\ \frac{dp_{BC}}{dt} &= q_{BB}p_{BC}(t) + q_{BC}p_{CC}(t) = -\frac{1}{7}p_{BC}(t) + \frac{9}{70}p_{CC}(t) \end{aligned}$$

Here $p_{CC}(t) = \exp(-t/14)$.

(viii) The boundary conditions are $p_{AC}(0) = p_{BC}(0) = 0$.

January 2005 Question 1

(i) Bookwork.

(ii) $X_i = -\mu \log U_i$.

January 2005 Question 5

(i) Bookwork

(ii)
   a) For any i, the holding time in state i, T_i, has exponential distribution with rate 3, so its variance is 1/9 per unit time. Holding times are independent, so the variance $T_i + T_{i+1}$ is $\mathrm{Var}\,(T_i) + \mathrm{Var}\,(T_{i+1}) = \frac{2}{9}$.
   b) N_3 is a Poisson random variable with mean 9, so has variance 9.

c) $\mathbb{E}(N_1 N_2) = \mathbb{E}(N_1^2) + \mathbb{E}(N_1(N_2 - N_1))$.
Now $\mathbb{E}(N_1^2) = \text{Var}(N_1) + [\mathbb{E}(N_1)]^2 = 3 + 9 = 12$.
And, by the independent increments property, N_1 and $N_2 - N_1$ are independent, so $\mathbb{E}(N_1(N_2 - N_1)) = \mathbb{E}(N_1)\,\mathbb{E}(N_2 - N_1) = 9$.
Adding these together we get $\mathbb{E}(N_1 N_2) = 21$.

(iii) a) The number of hits referred from Poodle is a thinned Poisson process, so is itself a Poisson process with parameter 5 per day. By the independent increments property, the number of hits on previous days is irrelevant. The probability we require is $1 - e^{-5}(1 + 5 + 5^2/2)$.

b) This probability is $e^{-2.5}$.

c) If Poodle's popularity is higher in some parts of the world than in others, there may be more hits on the website during some periods than during others, which might mean that a time-inhomogeneous PP might be a better model. I expect there are other reasons as well.

January 2005 Question 6

(i) The Markov property has a number of implications. One is that the time spent in state 1 has an exponential distribution (similarly for state 2). Another is that the time spent in state 2 is independent of the time spent in state 1. The first can be tested with a χ^2 goodness-of-fit test, the second with a scatter diagram and/or a contingency table.

(ii) a) The generator matrix is

$$Q = \begin{pmatrix} -0.2 & 0.2 & 0 \\ 0 & -0.5 & 0.5 \\ 0 & 0 & 0 \end{pmatrix}.$$

b) The condition is irrelevant. The probability is $e^{-0.6}$.

c) We need

$$\int_3^5 f_T(u) \times r_{12}\, du = e^{-0.6} - e^{-1}.$$

d) Now the probability is

$$\int_4^7 f_T(u)\left(1 - e^{-0.5(8-u)}\right) = e^{-0.8} - e^{-1.4} - 0.2e^{-4} \int_4^7 e^{0.3u}\, du = e^{-0.8} - e^{-1.4} - \frac{2}{3}(e^{-1.9} - e^{-2.8}).$$