Solutions to question for the CT - 4 (Part I) November 2005:

Solutions to Q1

(i) Level at the start of this year after:

Level at the start of the previous year	0 claims in the previous year	1 claim in the previous year	2 claims in the previous year	3 or more claims in the previous year
5	4	5	5	5
4	3	5	5	5
3	2	4	5	5
2	1	3	4	5
1	1	1	2	5

For each policyholder, the number of claims in each year has a Poisson (0.25) distribution. So

$$P(0\text{claims}) = e^{-0.25} = 0.7788$$

$$P(1\text{claims}) = 0.25e^{-0.25} = 0.1947$$

$$P(0\text{claims}) = (0.25)^2 \frac{e^{-0.25}}{2} = 0.0243$$

$$P(3\text{or more claims}) = 1 - 0.7788 - 0.1947 - 0.0243 = 0.0022$$

Thus, the transition matrix P is given by

$$\begin{pmatrix} 0.9735 & 0.0243 & 0 & 0 & 0.0022 \\ 0.7788 & 0 & 0.1947 & 0.0243 & 0.0022 \\ 0 & 0.7788 & 0 & 0.1947 & 0.0265 \\ 0 & 0 & 0.7788 & 0 & 0.2212 \\ 0 & 0 & 0 & 0.7788 & 0.2212 \end{pmatrix}$$

(ii) In order to be in level 1 in year 3, the policyholder requires two consecutive claim-free years. The probability of that is $90.7788)^2 = 0.6065$.
A similar argument can be used for the probability of being in level 3 in year 3, but a simpler argument might be to calculate the whole vector of probabilities x_3. Thus, we have

$x_1 = (00100)$
$x_2 = (00100).P = (00.778800.19470.0265)$
$x_3 = (00.778800.19470.0265).P = (0.606500.30330.03960.0506)$

Thus the probability of being in level 3 is 0.3033.

(iii)

a. The required conditions are that the chain is irreducible and aperiodic.

b. Irreducibility: level i can be reached from level j in $|j - i|$ steps; Aperiodicity: $p_{ii} > 0$ for some i

c. For the stationary distribution to be independent of the initial position, we need

$$(\pi_1 \pi_2 \pi_3 \pi_4)P = (\pi_1 \pi_2 \pi_3 \pi_4)$$

This implies

$$
\begin{aligned}
0.9735\pi_1 + 0.7788\pi_2 &= \pi_1 && (1)\\
0.0243\pi_1 + 0.7788\pi_3 &= \pi_2 && (2)\\
0.1947\pi_2 + 0.7788\pi_4 &= \pi_3 && (3)\\
0.0243\pi_2 + 0.1947\pi_3 + 0.7788\pi_3 &= \pi_4 && (4)\\
\text{and} \qquad \pi + 1 + \pi_2 + \pi_3 + \pi_4 &= 1 && (5)
\end{aligned}
$$

Solving these equations simultaneously we get

$$
\begin{aligned}
\pi_2 &= 0.0340\pi_1\\
\pi_3 &= 0.01244\pi_1\\
\pi_4 &= 0.00747\pi_1\\
\text{and} \qquad \pi_5 &= 0.00541\pi_1
\end{aligned}
$$

Thus, using these values and equation 5, we have

$$\pi_1 = 0.9440, \pi_2 = 0.0321, \pi_3 = 0.0117. \pi_4 = 0.0071 \qquad \text{and} \qquad \pi_5 = 0.0051.$$

(iv) A chi-square goodness of fit test is the most appropriate one.

Solutions to Q2

(i) State 0 implies not defeated last week; State 1 implies 1 defeat; State 2 implies 2 defeats in a row and State 3 implies Mr. Channel is fired (3 defeats in a row).

Thus, the transition probability matrix is

	State 0	State 1	State 2	State 3
State 0	0.8	0.2	0	0
State 1	0.8	0	0.2	0
State 2	0.8	0	0	0.2
State 3	0	0	0	1

(ii) P(k = 3) = P(3 defeats in a row) = $(0.2)^3) = 8 * 10^{-3}$

P(k = 4) = $0.8 * (0.2)^3) = 1.6 * 10^{-5}$

P(k = 5) = $0.8 * (0.2)^3) = 1.6 * 10^{-5}$

P(k = 6) = $0.8 * (0.2)^3) = 1.6 * 10^{-5}$

P(k = 7) = $\{0.8 * (0.2)^3\} * \{1 - (0.2^3)\})$.

(iii) Let e_i, i = 0, 1, 2 be the expected number of weeks until Mr. Channel gets fired. Starting from state i, clearly e_0 = E(k). Thus

$$
\begin{aligned}
e_0 &= 1 + 0.8e_0 + 0.2e_1\\
e_1 &= 1 + 0.8e_0 + 0.2e_2\\
e_2 &= 1 + 0.8e_0
\end{aligned}
$$

Thus,

$$e_0 = 5 + e_1 \Rightarrow e_1 = e_0 - 5 = 1 + 0.8e_0 + 0.2(1 + 0.8e_0)$$
$$\Rightarrow 0.04e_0 = 6.2$$
$$\Rightarrow e_0 = \frac{6.2}{0.04} = 155\text{weeks}.$$

(iv) The chain now is

$$\left(\begin{array}{c|ccc} & \text{State 0} & \text{State 1} & \text{State 2} \\ \hline \text{State 0} & 0.8 & 0.2 & 0 \\ \text{State 1} & 0.8 & 0 & 0.2 \\ \text{State 2} & 1 & 0 & 0 \end{array}\right)$$

Expected cost is $\$10,000\Pi_2$ where $(\Pi_1\Pi_2\Pi_3\Pi_4)$ is the stationary distribution. We thus have

$$0.8\Pi_0 + 0.8\Pi_1 + \Pi_2 = \Pi_0$$
$$0.2\Pi_0 = \Pi_1 0$$
$$0.2\Pi_1 = \Pi_2$$

Solving these equations, we get $\Pi_2 = \frac{1}{31}$. Thus the average amount paid in bribes per week by Mr. Channel is

$$\$\frac{10,000}{31}.$$

Solutions to Q3

(i) The generator matrix of the process would be

$$\left(\begin{array}{c|ccccc} & \text{State A} & \text{State F} & \text{State I} & \text{State O} & \text{State D} \\ \hline \text{State A} & -1 & 0.4 & 0.1 & 0.5 & 0 \\ \text{State F} & 0 & -\frac{1}{3} & \frac{1}{12} & \frac{1}{12} & \frac{1}{6} \\ \text{State I} & 0 & 0 & -\frac{1}{60} & 0 & \frac{1}{60} \\ \text{State O} & 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} \\ \text{State D} & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

(ii) The probability of ever visiting state I is $\frac{1}{10} + (\frac{4}{10} * \frac{1}{4}) = \frac{1}{5}$.

(iii)

a. $\frac{d}{dt}p_{AA}(t) = -p_{AA}(t)$, which has the solution $p_{AA}(t) = e^{-t}$.

b. Similarly, $\frac{d}{dt}p_{AF}(t) = -\frac{1}{3}p_{AF}(t) + 0.1p_{AA}(t)$, so that

$$\frac{d}{dt}\{e^{\frac{t}{3}}p_{AF}(t)\} = 0.1e^{\frac{t}{3}}p_{AA}(t) = 0.1e^{-\frac{2t}{3}}$$
$$\Rightarrow p_{AF}(t) = e^{-\frac{t}{3}} * 0.6(1 - e^{-\frac{2t}{3}}).$$

(iv)

a. The equation arises as follows: when the process is in state i, the subsequent holding time has mean λ_i, after which the process jumps to a different state, choosing state j with probability $p_{ij} = \frac{\sigma_{ij}}{\lambda_i}$ (independent of the length of the holding time. The total time to reach state D, is therefore the time until the first jump plus the time from arriving in the new state until hitting D (unless the new state is D).

b. We have $m_I = 60$, $m_O = 2$, $m_F = 3 + (\frac{1}{4} * 60) + (\frac{1}{4} * 2) = 18.5$, $m_A = 1 + (0.1 * 60) + (0.5 * 2) + (0.1 * 18.5) =$ 15.4 hours.

(v) The time - homogeneous Markov model has exponential holding times, so the distribution is completely determined by the expectation.

(vi) A simple check on whether the Markov model fits the data is to verify that the distributions of the holding times are roughly exponential, and a simple way of doing that is to compare sample standard deviations with the sample means. More detailed comparisons may be possible, depending on the size of the data set.

(vii)

a. Calculations required in the first case would include working out the expected duration of stay if the change were to be implemented, which involves solving the equation in part (iv) again. For the second situation, just replace m_O in the original calculations. New parameter values will need to be guessed. Whichever model comes out better should be compared with the initial situation, to deter mine whether the improvement was worth the additional resources.

b. As far as model suitability is concerned, on one hand the required decision is masked in terms of expectations, which lend themselves well to Markov process analysis. On the other, the fundamental problem in the system is queue length, which can never be successfully modeled by a process which tracks only a single individual at a time (probably a network of queuing processes would be a better model).