Cribs for the 3E3 Modelling Risk Exam Paper (2014-2015)

Q1(a)(i) The Markov Chain has four states: B (Birmingham), C (Cambridge), E (Exeter), and L (London). The transition probability matrix is

$$P = \begin{pmatrix} 0.7 & 0.06 & 0.18 & 0.06 \\ 0 & 0.7 & 0.18 & 0.12 \\ 0 & 0.15 & 0.7 & 0.15 \\ 0.03 & 0.03 & 0.24 & 0.7 \end{pmatrix}.$$

Q1(a)(ii) The initial state distribution is $q^0 = (0.25, 0.25, 0.25, 0.25)$. The state distribution after two weeks is

$$q^0 \times P \times P = (0.1825, 0.235, 0.325, 0.2575) \times P = (0.135475, 0.231925, 0.36445, 0.26815).$$

Q1(a)(iii) The steady state distribution is the solution of the following equations:

$$(u_1 \; u_2 \; u_3 \; u_4)P = (u_1 \; u_2 \; u_3 \; u_4).$$

The solution is

$$(u_1 \; u_2 \; u_3 \; u_4) = (0.031088461, 0.244226905, 0.413853129, 0.310831505).$$

Therefore, the numbers of cars in four locations are

$$400 \times (0.031088461, 0.244226905, 0.413853129, 0.310831505)$$

$$= (12.4, 97.7, 165.5, 124.3).$$

Therefore, there would be a long run space availability problem in two locations: Exeter and London.

Q1(a)(iv) We need to calculate the expected first passage times, which satisfy the following system of equations:

$$E(H_{ij}) = 1 + \sum_{k \neq j} E(H_{kj}) P_{ik}, \forall i.$$

For Cambridge, index $j = 2$. Let $G_{ij} = E(H_{ij})$.

$$\begin{aligned} G_{12} &= 1 + G_{12}P_{11} + G_{32}P_{13} + G_{42}P_{14} \\ G_{22} &= 1 + G_{12}P_{21} + G_{32}P_{23} + G_{42}P_{24} \\ G_{32} &= 1 + G_{12}P_{31} + G_{32}P_{33} + G_{42}P_{34} \\ G_{42} &= 1 + G_{12}P_{41} + G_{32}P_{43} + G_{42}P_{44}. \end{aligned}$$

The solution of the system is

$$G_{12} = 11.27, G_{22} = 4.09, G_{32} = 9.27, G_{42} = 11.88.$$

Therefore, the average number of weeks that elapse for a car that starts at Cambridge and return to Cambridge first time $G_{22} = 4.09$ weeks.

Q1(b) n: stage/week index

x_n: action taken in stage n, which is the number of labour force

s_n: state of the system in stage n where $s_n = x_{n-1}$

b_n: the minimum labour force in stage n

C_1: unit excess labour force cost in stage n

C_2: unit hiring cost in stage n

$s_{n+1} = g_n(s_n, x_n)$: state transition

$c_n(s_n, x_n)$: system cost in stage n. We have

$$c_n(s_n, x_n) = C_1(x_n - b_n) + C_2(x_n - s_n) = C_1(x_n - b_n) + C_2(x_n - x_{n-1}).$$

$f_n^*(s_n)$: optimal reward-to-go in stage n when state is s_n.

We have the optimality equations:

$$f_n^*(s_n) = \min_{x_n \geq b_n} \{C_1(x_n - b_n) + C_2(x_n - s_n) + f_{n+1}(x_n)\}, n = 1, \ldots, N$$

and

$$f_{N+1}^*(s_n) = 0.$$

Q1(c) We look at the following attributes in Summary Outputs when we assess the strength and weakness of a multiple regression model.

- R-square statistics. A larger R-square statistic implies that the regression model fits the data better.
- The standard error for the regression model. A smaller standard error gives a higher prediction power.
- The t-statistic (or p-value or confidence intervals) for the slope of each independent variable. In order for an independent variable to be significant in the regression model, a larger value of the t-statistic is preferred.
- We need to pay attention to multi-collinearity, which states that two independent variables are highly correlated and may give an incorrect impression that either of the two independent variables are not true drivers for the dependent variable.
- Other possible factors are: the error plot, the sign of the slope, sample size, other key drivers for the dependent variable, etc.

Q1. Examiner's Comment:

This very popular question is properly answered by most candidates. Most candidates did not have any problems with understanding and calculations of Markov chain. Majority of candidates had difficulties with understanding of the dynamic programming concept. Many candidates were able to correctly interpret regression analysis results.

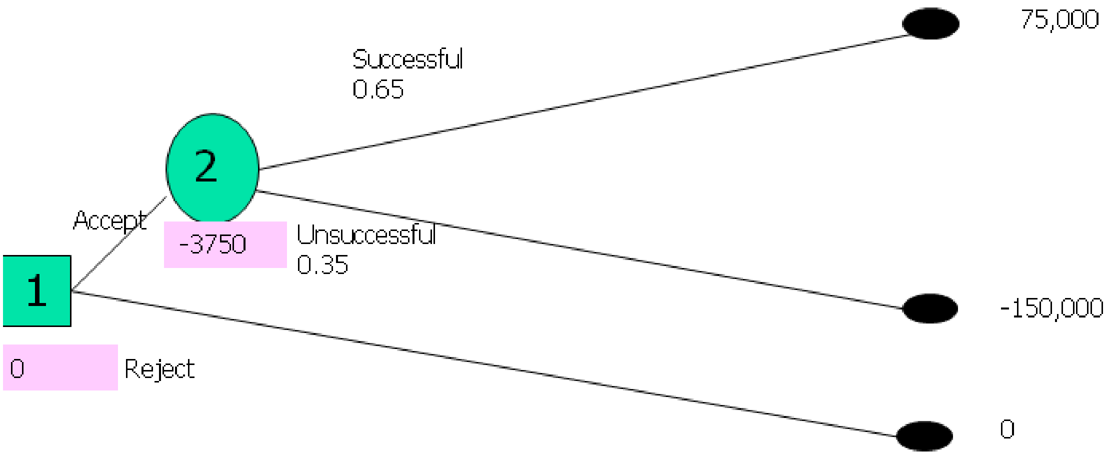

Figure 1: The decision tree without a review.

Q2(a)(i) See Figure 1.

Q2(a)(ii) The optimal decision is to reject the proposal. The expected profit is £0.

Q2(a)(iii) See Figure 2.

Q2(a)(iv) The optimal decision is to have a review first. If the review is favourable, then accept the proposal. Otherwise, reject the proposal. The expected profit is £13,125.

Q2(a)(v) If the cost for the manuscript review is £5,000, then it is still profitable for Business Publishing because 13,125 - 5,000 = £8,125. By checking the decision tree again, the maximum amount that Business Publishing would be willing to pay is £13,125.

Q2(b) For the $M/M/1$ queue, we have

$$c_n = \left(\frac{\lambda}{\mu}\right)^n = \rho^n.$$

Note that

$$\sum_{n=0}^{\infty} p_n = 1, \qquad p_n = \rho^n p_0.$$

The average number of people in the queue is given by

$$L = \sum_{i=0}^{\infty} i\rho^i (1-\rho).$$

This is just the mean of a geometric distribution which is $L = \frac{\rho}{1-\rho}$ or $\frac{\lambda}{\mu-\lambda}$.

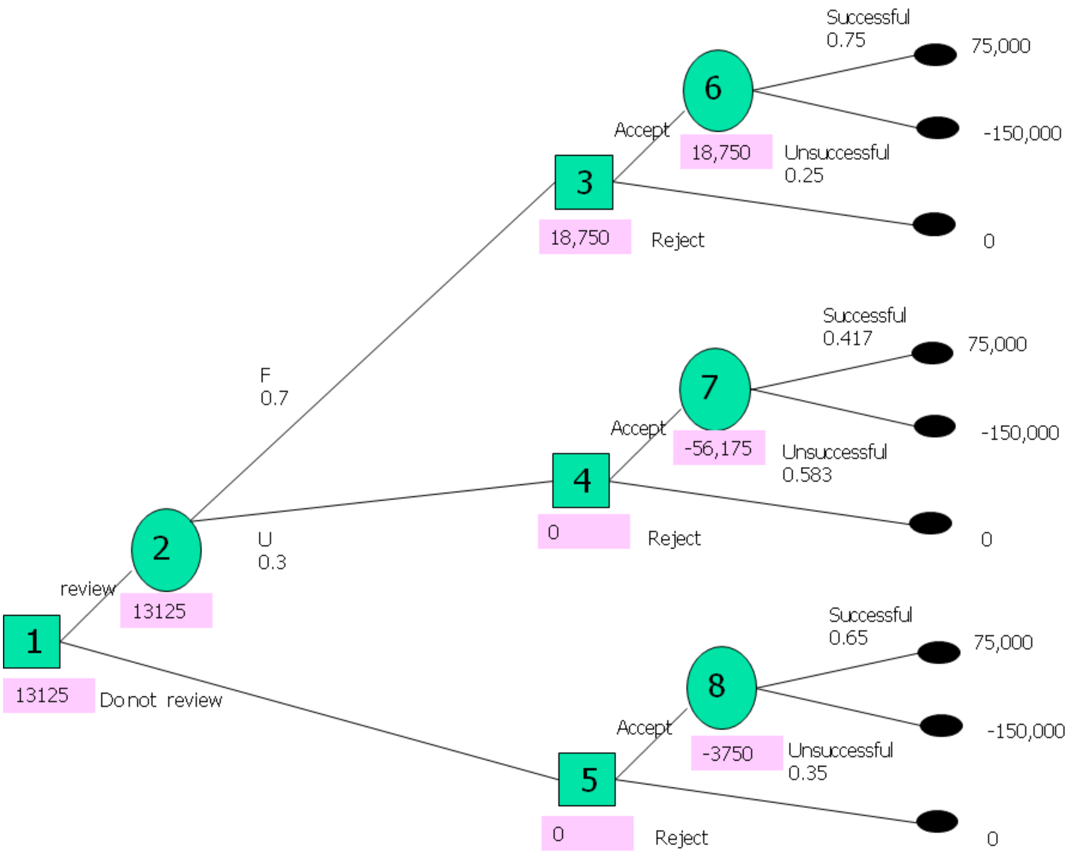

Figure 2: The decision tree with a review.

Let us now derive the formula for L.

$$\begin{aligned}
L &= \sum_{i=0}^{\infty} i\rho^i(1-\rho) \\
&= \rho\sum_{i=0}^{\infty} i\rho^{i-1} - \rho^2\sum_{i=0}^{\infty} i\rho^{i-1} \\
&= \rho\left(\sum_{i=0}^{\infty} \rho^i\right)' - \rho^2\left(\sum_{i=0}^{\infty} \rho^i\right)' \\
&= (\rho-\rho^2)\left(\tfrac{1}{1-\rho}\right)' \\
&= (\rho-\rho^2)\tfrac{1}{(1-\rho)^2} \\
&= \tfrac{\rho}{1-\rho}.
\end{aligned}$$

Given the formula for L, we can obtain formulae for the other performance measures by using Little's formula and the relation $W = W_q + 1/\mu$:

$$L_q = \frac{\lambda^2}{\mu(\mu-\lambda)}, \qquad W_q = \frac{\lambda}{\mu(\mu-\lambda)}, \qquad W = \frac{1}{\mu-\lambda}.$$

Q2(c) The Winters exponential smoothing method is defined by the following equations:

the base:

$$E_t = \alpha\frac{X_t}{S_{t-c}} + (1-\alpha)(E_{t-1} + T_{t-1}),$$

the trend:

$$T_t = \beta(E_t - E_{t-1}) + (1-\beta)T_{t-1},$$

and the seasonality (or season index):

$$S_t = \gamma\frac{X_t}{E_t} + (1-\gamma)S_{t-c},$$

where c is the number of time periods in the season. Because we have three updating equations, this method is often referred to as triple exponential smoothing.

The forecast is then given by

$$F_{t+k} = (E_t + kT_t)S_{t+k-c}.$$

The smoothing parameters are α, β, and γ all lie in the range 0 to 1.

For the simple exponential smoothing method, we define the smoothed average at time t to be

$$E_t = \alpha X_t + (1-\alpha)E_{t-1} = E_{t-1} + \alpha(X_t - E_{t-1}).$$

The (one-step) forecast is

$$F_{t+1} = E_t$$

The 2-step forecast is

$$F_{t+2} = F_{t+1} + \alpha(F_{t+1} - E_t) = F_{t+1}$$

and thus $F_{t+k} = F_{t+1}$ for all $k \geq 1$.

When the time series shows both trend and seasonal features, it is much better to use the Winters exponential smoothing method because the latter takes both trend and seasonal features into account.

Q2. Examiner's Comment:

This is also a very popular question and was very well answered in general. Most candidates had a very good understanding of the decision tree concept and associated calculations. Majority of the candidates understood queuing theory but some struggled to derive closed-form solutions. Finally, they well explained concepts of seasonality and trend in time series forecasting.

Q3(a)(i) Define H_{ij} to be the time of first passage from state i to state j and $f_{ij}(k)$ to be the probability that $H_{ij} = k$. Then by the definition we have

$$E(H_{ij}) = f_{ij}(1) + 2f_{ij}(2) + 3f_{ij}(3) + \dots$$

The law of total probability shows that

$$E(H_{ij}) = E(H_{ij}|C_1)P(C_1) + \dots + E(H_{ij}|C_n)P(C_n),$$

if $C_1, \dots, C_n$ are mutually exclusive and their union has a probability of one.

Now assume C_k to be the first transition going from state i to state k. Then we have

$$E(H_{ij}|C_j) = 1$$

because conditioning on the fact that the first transition goes from state i is state j, the expected number of steps from state i to state j is equal to one. Following the similar argument, we also have

$$E(H_{ij}|C_k) = 1 + E(H_{kj})$$

Therefore, we have for fixed j and any i

$$\begin{aligned} E(H_{ij}) &= E(H_{ij}|C_1)P(C_1) + \dots + E(H_{ij}|C_n)P(C_n) \\ &= \textstyle\sum_k E(H_{ij}|C_k)P(C_k) \\ &= \textstyle\sum_k E(H_{ij}|C_k)P_{ik} \\ &= P_{ij} + \textstyle\sum_{k\neq j} E(H_{ij}|C_k)P_{ik} \\ &= P_{ij} + \textstyle\sum_{k\neq j}(1 + E(H_{kj}))P_{ik} \\ &= 1 + \textstyle\sum_{k\neq j} E(H_{kj})P_{ik}. \end{aligned}$$

The above gives us a system of n linear equations with n unknowns $E(H_{ij})$.

Q3(b)(i) The arrival rate is $\lambda = 10 \times 1/7 = 10/7$ per hour and the service rate is $\mu = 1/0.5 = 2$ per hour.

Assume the number of servers $s = 1$. Then we have a $M/M/1$ queue.

Thus the average number of people in the queue is given by

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

Given the formula for L, we can obtain formulae for the other performance measures by using Little's formula and the relation $W = W_q + 1/\mu$:

$$L_q = \frac{\lambda^2}{\mu(\mu-\lambda)}, \qquad W_q = \frac{\lambda}{\mu(\mu-\lambda)}, \qquad W = \frac{1}{\mu-\lambda}.$$

Therefore,

$$L = 2.5,$$

$$W = 7/4,$$

$$W_q = 5/4,$$

$$L_q = 25/14.$$

Assume the number of servers $s = 2$. Then we have a $M/M/2$ queue. The average queue length is

$$L_q = \frac{(\lambda/\mu)^{s+1}}{(s-1)!(s-\lambda/\mu)^2} p_0,$$

where

$$p_0 = \frac{1}{\sum_{n=0}^{s-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^s}{s!}\left(\frac{s\mu}{s\mu-\lambda}\right)}.$$

Furthermore, we still use Little's formula and the relation $W = W_q + 1/\mu$:

$$W_q = \frac{L_q}{\lambda}, \qquad W = W_q + \frac{1}{\mu}, \qquad L_q = \lambda W_q.$$

Therefore,

$$L = 0.8187,$$

$$W = 0.5731,$$

$$W_q = 0.0731,$$

$$L_q = 0.1044.$$

Q3(c) The width L of the 95% confidence interval is approximately 4 STEM, where STEM $= \frac{\sigma}{\sqrt{n}}$.

When sample size $n = 100$, STEM $= \frac{\sigma}{10}$

If sample $n = 400$, STEM $= \frac{\sigma}{20}$.

If sample $n = 900$, STEM $= \frac{\sigma}{30}$.

If sample $n = 1600$, STEM $= \frac{\sigma}{40}$.

Figure 3 shows that STEM decreases in sample size n. However, the decrease slows down as sample increases. That is, the improvement of having a smaller 95% confidence interval becomes smaller when sample size n is larger, which is why many samples contain 500 to 1000 respondents.

Q3(d)(i) The R-square statistics for simple regression is defined by

$$\frac{\sum_{i=1}^{n}(a + bx_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2},$$

where n is the sample size of the data, x_i is the independent variable, y_i is the dependent variable, $\bar{y}$ is the average value of y_i, a and b are the y-intercept and the slope for the regression equation that is generated based on the least square method.

The R-square statistics represents the proportion of the variation in the dependent variable that can be explained by the regression equation. The R-square statistics indicates how the regression line fits the data and it is always between 0 and 1. If the R-square statistics is close to 1, then the regression line fits the data very well. If the R-square statistics is close to 0, then the regression line does not fit the data very well.

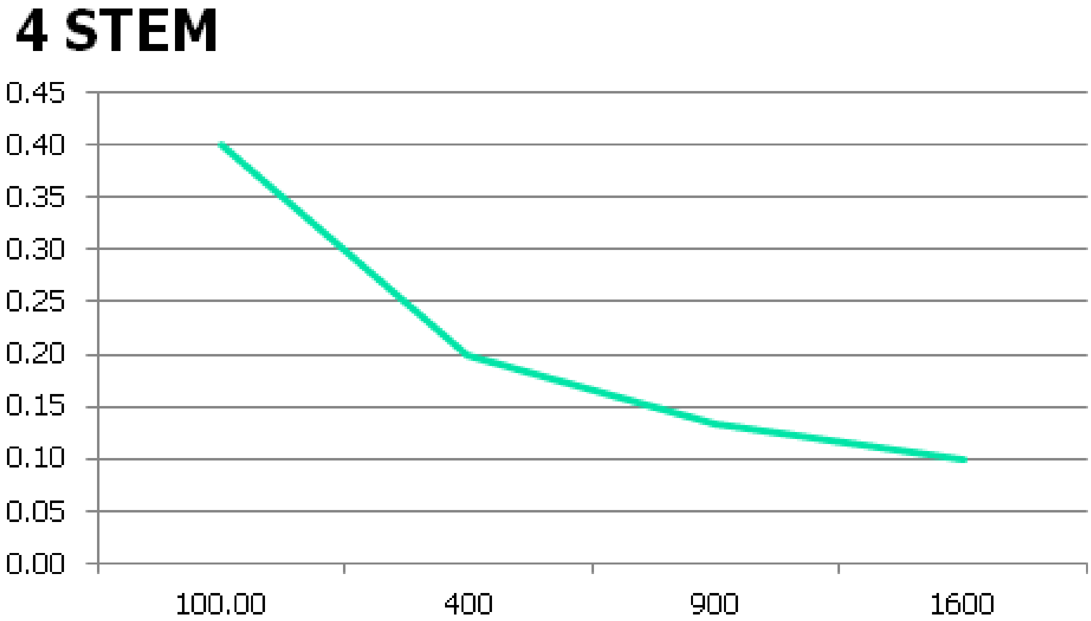

Figure 3: Sensitivity analysis of the width of the 95% confidence interval with respect to the sample size.

Q3(d)(ii) The simple linear regression assumes that the dependent variable y and the independent variable x have the following relationship:

$$y = \alpha + \beta x + \varepsilon,$$

where α is the y-intercept, β is the slope and ε represents the error term. Furthermore, it is assumed that ε is normally distributed with a mean equal to 0 and standard deviation equal to σ, where σ is independent of the value for independent variable x.

Q3(e) A number of portfolios generated from three stocks are shown in the risk-return scatter diagram. Three portfolios are highlighted called A, B and C in the diagram: A for the left most square, B for the middle square and C for the right most square.

Portfolio A is not efficient because there are portfolios that are better than A in both return and risk. Portfolios B and C are both efficient because we cannot identify any portfolio on the diagram that is better than B or C in both risk and return. Therefore portfolio A is not recommended, it is up to individual's risk appetite to choose between portfolio B and portfolio C as portfolio C has a higher return and higher risk compared with portfolio B.

Q3. Examiner's Comment:

This is an unpopular question. The candidates had difficulties in providing a correct mathematical proof for a Markov chain question. For queuing analysis they understood the overall concept, but many candidates were not able to apply more complicated queuing formulas for a simpler queue where the number of servers is one or two. In confidence intervals and the Excel output they understood the concept well but they struggled to understand sensitivity analysis of the regression results. They also had a good understanding of portfolio management.