# Cribs for the 3E3 Modelling Risk Exam Paper (2012-2013)

Q1(a)(i) For the $M/M/1$ queue,

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

Note this formulae can be derived from the queueing formulas in the data sheet. Therefore, the average number of customers in the new system is equal to the average number of customers in the old system if both arrival and service rates are trebled.

Q1(a)(ii) By Little's law, we have

$$W = \frac{L}{\lambda}.$$

Therefore, the the average time that a customer in the new system is equal to one third of the average time in the old system if both arrival and service rates are trebled.

Q1(a)(iii) By the queueing formulas in the data sheet, we have

$$p_0 = \frac{1}{1 + \frac{\lambda}{\mu}\frac{\mu}{\mu-\lambda}} = 1 - \frac{\lambda}{\mu}.$$

$$p_n = (\frac{\lambda}{\mu})^n p_0, \forall n \geq 1.$$

Therefore, the steady-state probability distribution does not change if both arrival and service rates are trebled.

Q1(b)(i) Let $M$ denote S&P 500, $P$ be the portfolio, $R = -0.3$ be the correlation coefficient between $M$ and $F$, and $\alpha = 0.5 = 50\%$. Note that $r_M = 0.1$, $\sigma_M = 0.04$, $r_F = 0.09$, $\sigma_F = 0.07$. Then the mean and variance of the portfolio $P$ are

$$r_P = \alpha r_M + (1 - \alpha) r_F = 0.095 = 9.5\%,$$

$$\begin{aligned} \sigma_P^2 &= \alpha^2 \sigma_M^2 + (1-\alpha)^2 \sigma_F^2 + 2\alpha(1-\alpha) R \sigma_M \sigma_F \\ &= 0.001205. \end{aligned}$$

Hence $\sigma_P = 0.034713$.

Q1(b)(ii) Let $M$ denote S&P 500, $P$ be the portfolio, $R = 0$ be the correlation coefficient between $M$ and $F$, and $\alpha = 1$ and $\beta = -1$. Let $D = F - D$. Note that $r_M = 0.1$, $\sigma_M = 0.04$, $r_F = 0.09$, $\sigma_F = 0.07$. Then the mean and variance of random variable $D$ are

$$r_D = \alpha r_M + \beta r_F = 0.01 = 1.0\%,$$

$$\begin{aligned} \sigma_D^2 &= \alpha^2\sigma_M^2 + \beta^2\sigma_F^2 + 2\alpha\beta R\sigma_M\sigma_F \\ &= 0.0065. \end{aligned}$$

Hence $\sigma_D = 0.080623$.

$$\begin{aligned} P(D > 0) &= P(\tfrac{D-r_D}{\sigma_D} > \tfrac{0-r_D}{\sigma_D}) \\ &= P(Z > 0.124035) \\ &= 0.450644. \end{aligned}$$

Or 0.4522 approximately.

Q1(c)(i) The correlation coefficient between two variables $X$ and $Y$ represents the strength of the linear relationship between these two variables. It is a number between -1 and 1. If it is close to 1, it means that $X$ and $Y$ are strongly and positively correlated and if one variable goes up, then it is most likely that another variable goes up. If it is close to -1, it means that $X$ and $Y$ are strongly and negatively correlated and if one variable goes up, then it is most likely that another variable goes down. If it is close to 0, it means that $X$ and $Y$ are not strongly correlated.

Q1(c)(ii) The $R$-square statistic is the proportion of the variation in the dependent variable that can be explained by the regression equation. It is a number between 0 and 1. If it is close to 1, then the regression equation fits the data very well. If it is close to 0, then the regression equation does not fit the data well.

Q1(c)(ii) The slope of a simple regression equation represents the change of the dependent variable for every unit change of the independent variable. The slope of a multiple regression equation has the similar meaning except for one condition which states that all other variables are assumed to be fixed.

Q2(a)(i) Because the demand in a period is either 1 or 2 or 3, and the inventory at the beginning of a period is at most 4, the inventory level at the end of a period is less than or equal to 3. Hence the probability, that the inventory level at the end of a period is 4, is zero.

Q2(a)(ii) The Markov Chain has four states: 0, 1, 2, 3, representing the inventory level at the end of a period. The transition probability matrix is

$$\begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}.$$

Q2(a)(iii) We need to find the stationary probability distribution $(u_0, u_1, u_2, u_3)$, which satisfies the following equations:

$$(u_0, u_1, u_2, u_3) \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix} = (u_0, u_1, u_2, u_3),$$

and

$$u_0 + u_1 + u_2 + u_3 = 1.$$

The solution of the equations is the following:

$$(u_0, u_1, u_2, u_3) = (11/48, 16/48, 12/48, 9/48).$$

Q2(a)(iv) The ordering numbers and costs are listed in the table below:

| state | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| order quantity | 4 | 3 | 0 | 0 |
| ordering cost | $5 + 4 \times 0.5$ | $5 + 3 \times 0.5$ | 0 | 0 |

The average total ordering cost in each period is

$$(5 + 4 \times 0.5) \times 11/48 + (5 + 3 \times 0.5) \times 16/48 = 77/48 + 104/48 = 181/48.$$

The average total inventory cost in each period is

$$2 \times (0 \times 11/48 + 1 \times 16/48 + 2 \times 12/48 + 3 \times 9/48) = 134/48.$$

Not meeting the demand occurs only when the inventory level is equal to 0 at the end of a period, which is the case where the inventory level at the beginning of the previous period is equal to either 2 or 3. When the inventory level at the beginning of the previous period is equal to 3, all demand is satisfied. When the inventory level at the beginning of the previous period is equal to 2, one unit demand is not met if demand in that period is equal to 3. Therefore, the average total penalty for not meeting the demand is

$$12/48 \times 1/3 \times 3 = 12/48.$$

Q2(b) Collect statistical data.

$$\begin{aligned} &n_1 = 35, n_2 = 35 \\ &m_1 = 288, m_2 = 279 \\ &\sigma_1 = 16.23, \sigma_2 = 15.91 \\ &\alpha = 5\%. \end{aligned}$$

Step 1. Make hypotheses. NH: $m_1 = m_2$. AH: $m_1 \neq m_2$.

Step 2. Calculate the $t$-statistic:

$$t = \frac{m_1 - m_2}{\text{STEDM}} = \frac{m_1 - m_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = 2.342740303,$$

where STEDM $= 3.841655$.

Step 3. Make a comparison. The absolute value of 2.342740303 is larger than 1.96, the $z$-value for the 5% significance level.

Step 4. Make a conclusion. At the 5% significant level, we think that the new system is different from the old system in terms of the average number of pages an ink cartridge can print.

Q2(c) A single smoothing exponential method is a special case of Winters' multiplicative smoothing method. The former has only one base smoothing parameter. The latter has three smoothing parameters for base, trend and seasonality.

It is fine if you want to explain (1) what trend and seasonality mean, and (2) why is it important to include trend and seasonality.

The former has one smoothing equation and the latter has three smoothing equations.

The prediction equations for two methods are also different.

Q3(a)(i) The regression equation is

$$Y = 25.344 + 0.086X_1 + 0.156X_2.$$

Explain all three variables $X_1$, $X_2$ and $Y$. Explain the intercept and two slopes. ...

Q3(a)(ii) When $X_1 = 1000$ and $X_2 = 36$, $Y = 25.344 + 0.086X_1 + 0.156X_2 = 116.96$. A 95% confidence interval is $[Y - 1.96 \times S_e, Y + 1.96 \times S_e] = [101.7563, 132.1637]$, where $S_e = 7.757$ is the standard error of the regression model. Therefore, with the 95% confidence level, we are pretty sure that the range of bookings for the hotel would be between 101 and 132.

Q3(a)(iii) The slope for 'Number of quotes' is 0.086. It means that on average the number of bookings is increased by 0.086 units for any additional quote. Hence, if the 'Number of quotes' is increased by 10, then the average increase for the number of bookings is 0.86. It is even better to use a range of values for the change of the number of bookings for an increase of 10 quotes.

Q3(a)(iv) We look at the following parameters and aspects:

1. $R$-square statistic,
2. Standard error for the regression equation,

3. $t$-statistic (or $p$-value) for each independent variable.

It seems that $t$-statistic for "Competitive price" is too small and smaller than 1.96, which may imply that this variable is not statistically significant. It may be caused by multi-collinearity because we think that two independent variables used in the regression equation are highly correlated. Therefore, it is worthwhile to remove one of these two variables and to check the new regression equation.

Q3(b)(i) The arrival rate is $\lambda = 1/10$ per minute and the service rate is $\mu = 1/3$ per minute. By the queueing formulas for the $M/M/1$ queque in the data sheet, we have

$$p_0 = \frac{1}{1 + \frac{\lambda}{\mu}\frac{\mu}{\mu-\lambda}} = 1 - \frac{\lambda}{\mu} = 0.7.$$

Therefore, the waiting probability is $1 - p_0 = 0.3$.

Q3(b)(ii) By the queueing formulas for the $M/M/1$ queque in the data sheet, we have

$$L_q = p_0 \times (\lambda/\mu)^2/(1-\lambda/\mu)^2 = \frac{\lambda^2}{\mu(\mu-\lambda)} = 0.13.$$

Q3(b)(iii) By Little's law, we have

$$W_q = \frac{L_q}{\lambda} = \frac{\lambda}{\mu(\mu-\lambda)}.$$

The service rate is fixed, in order for the average waiting time to be more than 3 minutes, we have

$$\frac{\lambda}{\mu(\mu-\lambda)} > 3,$$

which shows that $\lambda = 1/6$, i.e., the average interval arrival time is 6 minutes.

Q3(c) In the Micro-chip project, it is assumed that the demand is uncertain, the capacity is fixed, the unit price is fixed, the unit cost is fixed, and the investment cost is fixed. Suppose that the average demand is equal to the capacity. Then the value calculated based on the average demand is most likely not equal to the average value of the project. This is example for the flaw of averages because the upside opportunity (high demand) cannot balance the downside risk (low demand).

Q3(d) The what-if (sensitivity) analysis is very useful because it can help to identify how sensitive the project value is to a given uncertain parameter and it can help to identify the most important uncertain parameters that should be included in Monte Carlo Simulation. The drawbacks of sensitivity analysis are: (1) It does not take probabilities into account, (2) It can only address one uncertain parameter each time, (3) it does not take the correlation between uncertain parameters into account.

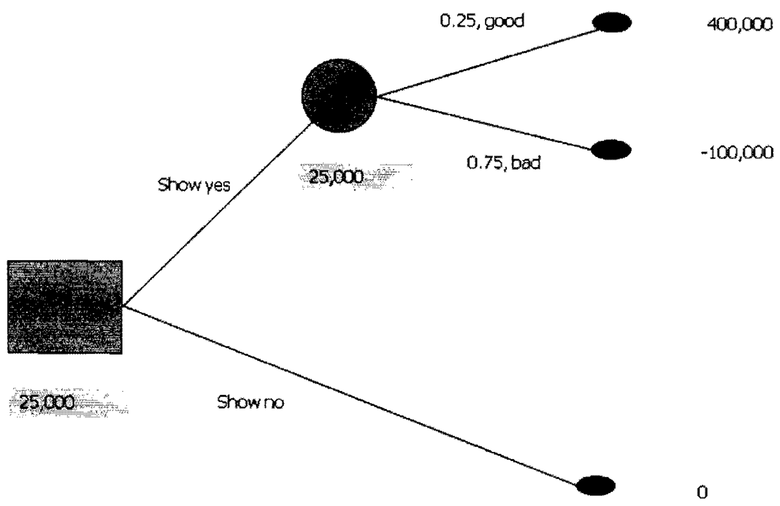


Figure 1: The decision tree without the market research.

Q4(a)(i) The decision tree without the market research is shown in Figure 1. The calculations are done in the tree. The EMV is \$25,000. The optimal decision is to televise all shows.

Q4(a)(ii) The decision tree without the market research when the prefect information is available is shown in Figure 2. The calculations are done in the tree. The optimal decisions are to televise shows when they are good and not to televise shows when they are bad. The EVPI = 100,000 - 25,000 = \$ 75,000.

Q4(a)(iii) It is necessary to calculate some conditional probabilities. Let $g$ and $b$ be the random variables that the market research firm will predict that the show to be good and bad, respectively. Let $G$ and $B$ be the random variables that shows are good and bad, respectively. Then the data shows that $P(G) = 0.25$, $P(B) = 0.75$, $P(g|G) = 0.9$, and $P(b|B) = 0.8$. By the total probability law, we have

$$P(g) = P(g|G)P(G) + P(g|B)P(B) = 0.375,$$

and $P(b) = 1 - P(g) = 0.625$. By Bayes' Theorem, we have

$$P(G|g) = \frac{P(g|G)P(G)}{P(g)} = 0.6, \qquad P(G|b) = \frac{P(b|G)P(G)}{P(b)} = 0.04,$$

$P(B|g) = 1 - P(G|g) = 0.4$, and $P(B|b) = 1 - P(G|b) = 0.96$.

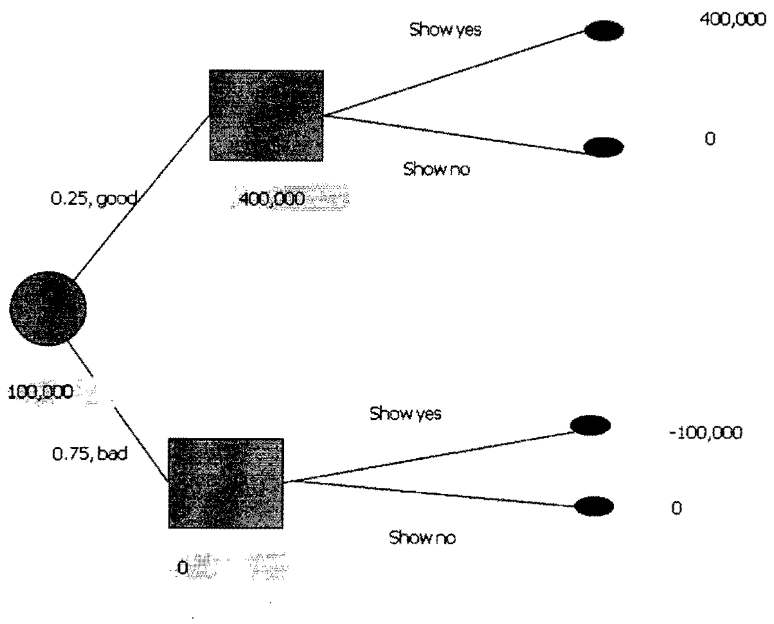


Figure 2: The decision tree without the market research when the prefect information is available.

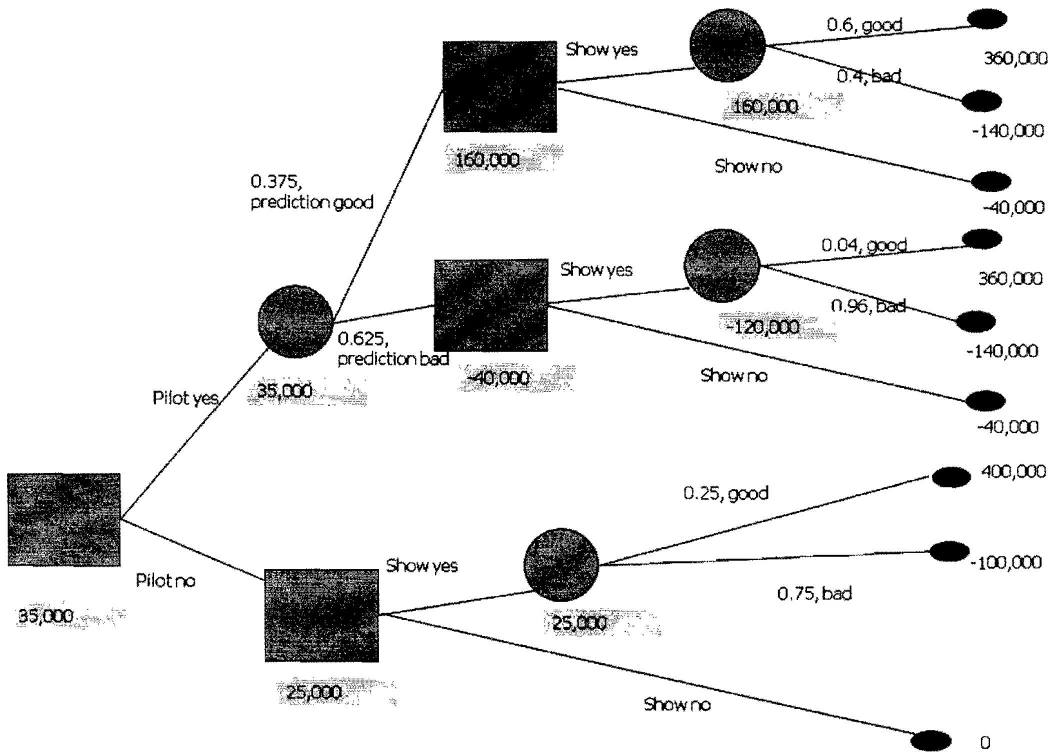


Figure 3: The decision tree with the market research.

The decision tree with the market research is shown in Figure 3. The calculations are done in the tree. The EMV with the sample information is \$35,000.

The EVSI = 35,000 - 25,000 = \$ 10,000.

Q4(b)(i) The Markov Chain has three states: I, II, III, representing the store index. The transition probability matrix is

$$\begin{pmatrix} 0.90 & 0.10 & 0 \\ 0.85 & 0.05 & 0.1 \\ 0.50 & 0.10 & 0.40 \end{pmatrix}.$$

Q4(b)(ii) The proportions of customers three stores will retain by February 1 are

$$(1/4, 1/3, 5/12) \begin{pmatrix} 0.90 & 0.10 & 0 \\ 0.85 & 0.05 & 0.1 \\ 0.50 & 0.10 & 0.40 \end{pmatrix} = (8.6/12, 1/12, 2.4/12).$$

The proportions of customers three stores will retain by March 1 are

$$(8.6/12, 1/12, 2.4/12) \begin{pmatrix} 0.90 & 0.10 & 0 \\ 0.85 & 0.05 & 0.1 \\ 0.50 & 0.10 & 0.40 \end{pmatrix} = (0.8158333, 0.09583333, 0.0883333).$$

Q4(b)(iii) The stationary probability distribution $(u_1, u_2, u_3)$ satisfies the following equations

$$(u_1, u_2, u_3) \begin{pmatrix} 0.90 & 0.10 & 0 \\ 0.85 & 0.05 & 0.1 \\ 0.50 & 0.10 & 0.40 \end{pmatrix} = (u_1, u_2, u_3),$$

and

$$u_1 + u_2 + u_3 = 1.$$

The solution is $(u_1, u_2, u_3) = (0.8888, 0.0952, 0.0158)$.

Q4(c)(i) The efficient frontier in the mean-variance diagram consists of a number of portfolios that are efficient. A portfolio $X$ is efficient if it does not exist another portfolio which has a better mean and a better variance than this portfolio $X$.

Q4(c)(ii) The meaning of the 95% confidence interval for the population proportion is the following: the 95% of all 95% confidence intervals for the population proportion contain the population proportion and 5% of all 95% confidence intervals for the population proportion do not contain the population proportion.

**ENGINERING TRIPOS PART IIA 2012-2013**

**ASSESSOR'S REPORT: MODULE 3E3 Modelling Risk**

**Overall:** Assessment performance was excellent and beyond the expectation of the lecturer. Only one student out of 95 had a raw mark below 40%. The average raw mark for the test is 71.95% with the maximum of 97.5% and the minimum of 29%.

The target distribution for the 94 candidates who had completed Part IB of the Engineering Tripos according to the Woodhouse scaling algorithm was as follows:

| | |
|---|---|
| I | 28-38 |
| IIi | 33-43 |
| IIii | 15-25 |
| III and below | 0-9 |

The raw marks were scaled based on the method provided in the template Excel file.

The distribution of the scaled marks was as follows (excluding one student who did not take Part IB):

| | |
|---|---|
| I | 36 |
| IIi | 34 |
| IIii | 18 |
| III and below | 6 |

**Q1 Queuing theory, portfolio management, regression analysis**

29 attempts. Average mark is 70.4% and standard deviation is 13.7%. Students often had more difficulties with qualitative questions in part (c) such as the meanings of R-square statistics and the slope of a regression equation.

**Q2 Markov chains, hypothesis testing, time series forecasting**

23 attempts. Average mark is 64.4% and standard deviation is 17.9%. This is the least popular question. Some students used an incorrect method for calculating inventory levels at the end of time periods. Most students did not make a great comparison between Winter's smoothing methos and the single-smoothing method.

**Q3 Regression analysis, queuing theory, Monte Carlo simulation**

56 attempts. Average mark is 63.7% and standard deviation is 17.5%. While most students indicated that the regression formula was overall good with possible drawbacks, the explanations were not very good. Students should have mentioned about R-square statistic, t-statistic, p-value, multi-collinearity etc. The qualitative question is part (c) was not well answered.

**Q4 Decision tree analysis, Markov chains**

82 attempts. Average mark is 79.7% and standard deviation is 13.7%. This is the most popular question and the raw average is very high. The main reason they did so well is because most students

did not make a lot of mistakes in part (a). Determining the value of sample information was challenging to some students. The concepts of efficient frontier and a 95% confidence interval for the population proportion were not well explained by some students.

| |
|---|
| Houyuan Jiang<br>05 June 2013 |