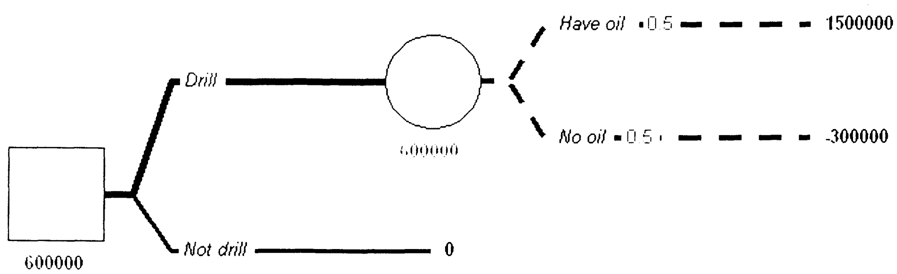
PART IIA 2005
3E3: Modelling risk
Principal Assessor: Dr G. Giallombardo JIMS

Datasheet: None

1

1(a) (i) The simple decision tree without survey information for the company is shown below.

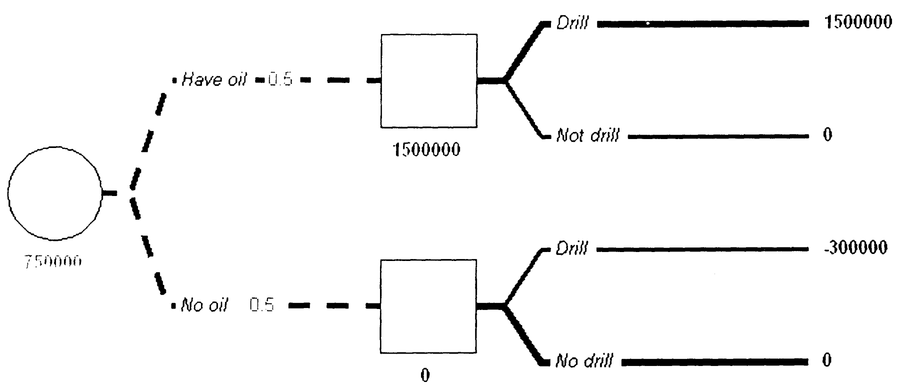


From the above decision tree, there are three possibilities:

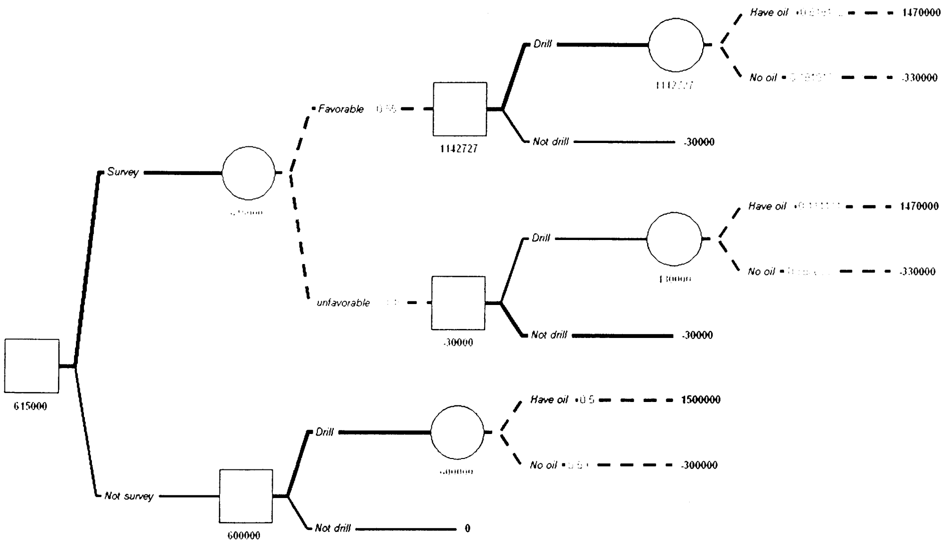
- Drill and Have oil: The payoff is 1,800,000 – 30,000 = $1,500,000.
- Drill and No oil: The payoff is – $30,000.
- Do not drill: The payoff is $0.

We now use the rolling back technique to find out the optimal strategy: At a chance node, we calculate the expected value of various scenarios, and at a decision node, we take the maximum value of various alternatives. The expected payoff at the top chance node is 0.5*1,500,000 + 0.5*(-30,000) = $600,000. The maximum value at the only decision node (root node) is $600,000, which is the maximum of 600,000 and 0. The expected monetary value for this decision problem is $600,000. Therefore, the optimal strategy for the company is to drill this site if no survey information is available.

1(a)(ii) The expected value of perfect information is equal to the optimal expected monetary value with perfect information minus the expected monetary value without perfect information. We can calculate the expected monetary value with perfect information by drawing another decision tree by swapping the order of the chance node and the decision node of the decision tree in (a). The new decision tree is shown below. The expected monetary value with perfect information is $750,000, which can be calculated using the rolling back technique. Therefore the EVPI is $150,000.



1(a)(iii) The decision tree with survey information included in shown below.



The expected monetary value for the decision tree with survey information included can be calculated using the rolling back technique again. A crucial step is to calculate a number of probability values at several chance nodes in this new decision tree. This can be done using the given conditional probabilities and Bayes' rule. Define a few stochastic events: A = "The site contains oil", B = "The site does not contain oil", X = "The survey indicates there is oil at the site", Y = "The survey indicates there is no oil at the site". It is known that P(X given A) = 0.9, P(Y given A) = 0.1, P(X given B) = 0.2 and P(Y given B) = 0.8. According to the company's estimate, P(A) = P(B) = 0.5.

Hence we obtain the following joint probability values: P(X and A) = P(X given A)*P(A) = 0.45, P(X and B) = 0.1, P(Y and A) = 0.05, P(Y and B) = 0.4. This implies that P(X) = P(X and A) + P(X and B) = 0.45 + 0.1 = 0.55, and P(Y) = P(Y and A) + P(Y and B) = 0.45.

The following conditional probabilities can be obtained using Bayes' rule again. P(A given X) = P(A and X)/P(X) = 0.45/0.55 = 0.818181, P(B given X) = 0.111111, P(A given Y) = P(A and Y)/P(Y) = 0.05/0.45 = 0.1818118, P(B given Y) = 0.888889. These probabilities are used in the new decision tree above.

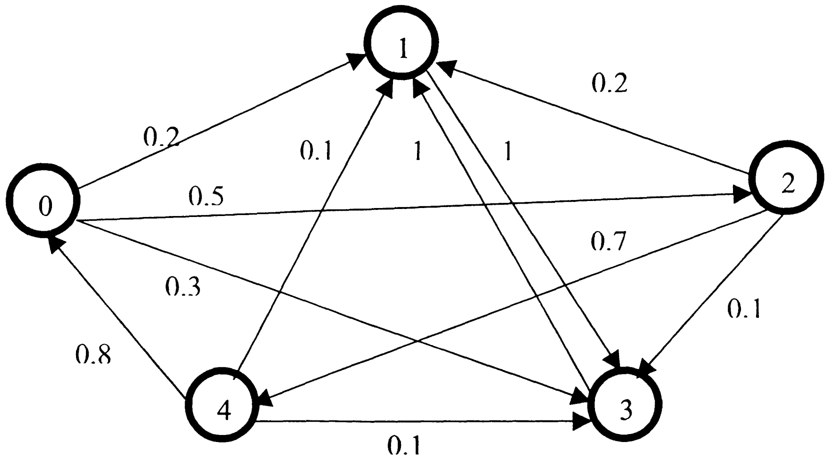
1(a)(iv) The expected value of sample information (EVSI) is equal to the expected money value with sample information minus the expected monetary value without sample information. Applying the rolling back technique again, the expected monetary value is \$615,000. Hence EVSI is 615,000 - 600,000 = \$15,000.

1(b) The Central Limit Theorem states that if we draw n samples independently from any probability distribution, then the sample mean is a random variable and it looks like a Normal distribution $N(\boldsymbol{m}, \boldsymbol{s}/n^{1/2})$, where $\boldsymbol{m}$ is the true population mean and $\boldsymbol{s}$ is the true population standard deviation.

As uncertain numbers are averaged together, uncertainty is reduced. This is known as diversification and is an important manifestation of the Central Limit Theorem. A business explanation of the Central Limit Theorem is to diversify your investments in unrelated areas so that your risk of losing money is reduced and you still receive the same average return. Suppose there are two unrelated investments on the market that have the same expected return and the same risk in terms of standard deviation. Then you will have a smaller chance to lose money by investing your money in both investments than in only one investment. On average, the more diversified you are, the less risk you will have of losing money. Here is a mathematical example. Suppose X_1 and X_2 are random variables representing throwing two dice respectively and $X = (X_1 + X_2)/2$. Then X has the same mean as X_1, but X has a smaller standard deviation than X_1.

The underlying assumption behind the Central Limit Theorem is that samples must be drawn independently from identically distributed random variables. If you diversify your investments in some highly correlated areas, you may not be able to reduce your risk by much.

3(a)(i) The transition network of this Markov chain is shown below.



States 1 and 3 are accessible from each other ($p_{31} = 1$ and $p_{13} = 1$), but no other states are accessible from these states ($p_{1j} = 0$ and $p_{3j} = 0$ for j = 0, 2, 4). Therefore, states 1 and 3 communicate and form one class of the Markov chain.

State 0 is accessible from state 4 ($p_{40} = 0.8$), state 2 is accessible from state 0 ($p_{02} = 0.5$), and state 4 is accessible from state 2 ($p_{24} = 0.7$), so each of these states is accessible from each of these other states. Therefore, states 0, 2, and 4 communicate and form a second class of the Markov chain.

3(a)(ii) Periodicity is a class property. State 1 can get back to itself in an even and only even number of transitions: 1 → 3 → 1. Therefore, the period of state 1 is 2 and so is state 3. Similarly, state 0 can get back to itself in three steps via the path 1 → 2 → 4 → 0 and state 0 can get back to itself with only a number of steps that is a multiple of three. Therefore, the period of state 0 is 3, and so are states 2 and 4. Another way to determine the period of each state is through use of the *n*-step transition probability matrix. Since $p_{11} = p_{33} = 0$ and $p^{(2)}_{11} = p^{(2)}_{33} = 1$, the class {1, 3} has period 2. Now note that $p_{00} = p_{22} = p_{44} = 0$, $p^{(2)}_{00} = p^{(2)}_{22} = p^{(2)}_{44} = 0$, and $p^{(3)}_{00} > 0$, $p^{(3)}_{22} > 0$, $p^{(3)}_{44} > 0$. This indicates that the class {0, 2, 4} has period 3.

3(a)(iii) This Markov chain is neither irreducible nor aperiodic as it as any irreducible (or aperiodic) Markov chain has all states in one class. This Markov chain is not regular as any regular Markov chain must be irreducible.

3(b)(i) Since P is regular as stated, this Markov chain can reach a steady state and the steady state probability distribution π must satisfy the following system of equations

$$\pi = \pi \mathrm{P}, \text{ and } \pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$$

The solution of this system of equations is $\pi = (16/99, 36/99, 31/99, 16/99)$.

3(b)(ii) The expected payoff of the Markov chain is equal to

$$C_1\pi_1 + C_2\pi_2 + C_3\pi_3 + C_4\pi_4 = 0*16/99 + 1000*36/99 + 3000*31/99 + 6000*16/99 = \$2272.727$$

3(b)(iii) Let X_{ij} = time of first passage from state i to state j and B_k denote the event of the first transition going from state i to state k. Then the expected first passage time from i to j is defined by

$$E(X_{ij}) = f_{ij}(1) + 2f_{ij}(2) + 3f_{ij}(3) + \ldots.$$

Where f_{ij} is the probability distribution of X_{ij}. Recall the conditional expectation formula

$$E(X_{ij}) = E(X_{ij}|B_1)P(B_1) + \ldots + E(X_{ij}|B_n)P(B_n)$$

Notice that

$$E(X_{ij}\,|B_j) = 1 \text{ and } E(X_{ij}|B_k) = 1+E(X_{kj}) \text{ when } k \neq j$$

It follows that

$$\begin{aligned} E(X_{ij}) &= \sum_k E(X_{ij} \mid B_k) P_{ik} \\ &= P_{ij} + \sum_{k \neq j} (1 + E(X_{kj})) P_{ik} \\ &= \sum_k P_{ik} + \sum_{k \neq j} E(X_{kj}) P_{ik} \\ &= 1 + \sum_{k \neq j} E(X_{kj}) P_{ik} \end{aligned}$$

Fixing j, we obtain a system of n linear equations with n unknowns $E(X_{1j})$, $E(X_{2j})$, ..., $E(X_{nj})$. Since the transition matrix P is given, $E(X_{ij})$ can be found by solving the above system of linear equations. Let i = 1, 2, 3 and 4 denote states New, OK, Worn and Fail and y_i be $E(X_{i1})$ respectively. We need to calculate y_1. Then we have the following equations:

$$y_1 = 1 + y_2 P_{12} + y_3 P_{13} + y_4 P_{14}$$
$$y_2 = 1 + y_2 P_{22} + y_3 P_{23} + y_4 P_{24}$$
$$y_3 = 1 + y_2 P_{32} + y_3 P_{33} + y_4 P_{34}$$
$$y_4 = 1 + y_2 P_{42} + y_3 P_{43} + y_4 P_{44}$$

The solution of the above system of linear equations is y = (495/80, 43/8, 7/2, 1). Therefore the expected first passage time from New to New is 495/80 $\cong$ 6.19 days.

2(a) Let N be the total number of customers served by this server in the sample and T be the total number of hours the server spent. Then the mean service time is equal to T/N. From the table, $T = 26$ and $N = 122$. Therefore the mean service time is equal to $T/N = 26/122 = 0.2131$ hours = 12.787 minutes and the service rate is equal to $N/T = 122/26 = 4.692$ customers per hour.

2(b) Little's formula states that $L = \lambda W$ and $L_q = \lambda W_q$, where λ, L, L_q, W, and W_q represent the customer arrival rate, the expected number of customers in the queueing system, the expected number of customers in the queue, the expected time spent in the system per customer, and the expected waiting time in the queue. The expected queue length is L_q when a new customer joins the queue. When he starts to receive service, the time that he spends in the system so far is W_q, which is his waiting time in the queue. During this waiting time, more customers join the queue. The expected number of new customers arriving during his waiting time is equal to λW_q. On the other hand, when he starts to receive his service, the expected number of customers in the queue is still L_q in a stationary queueing system. Therefore $L_q = \lambda W_q$. Similarly, we can derive $L = \lambda W$.

2(c)(i) The arrival rate is $\lambda = 24$ per hour, and the service rate is $\mu = ½$ per minute or 30 per hour. The traffic intensity is $\rho = 24/30 = 0.8$. Therefore, in a M/M/1 system, the average number of customers in the system is equal to $L = \rho/(1-\rho) = \lambda/(\mu-\lambda) = 4$ customers.

2(c)(ii) The average time that a customer spends in the system is equal to $W = L/\lambda = 1/(\mu - \lambda) = 1/(30-24) = 1/6$ hour = 10 minutes.

2(c)(iii) The fraction of the time that the system is idle is equal to $\rho^0 (1-\rho) = 1 (1-0.8) = 0.2$. The system is idle 20% of the time.

2(d)(i) In a birth and death process, the balance equations state that $E_n = L_n$, where E_n and L_n are the rate (average number of events per unit time) at which the system enters and leaves state *n*, respectively. Let p_n be the steady state probability that the birth and death process is in state *n*. The balance equation at state 0 is equivalent to

$$E_0 = \mu_1 p_1 = L_0 = \lambda_0 p_0$$

as the process can only go to state 1 from state 0 in one step with a birth rate λ_0 and can enter state 0 only from state 1 in one step with a death rate μ_1. Similarly the balance equation at state *n* is equivalent to

$$E_n = p_{n-1}\lambda_{n-1} + p_{n+1}\mu_{n+1}$$

$$L_n = p_n(\lambda_n + \mu_n)$$

as the process can go to state *n* from only states *n-1* and *n+ 1* in one step with a birth rate λ_{n-1} and a death rate μ_{n+1}, and the process can only go to states *n-1* and *n+1* from state *n* in one step with a death rate μ_n and λ_n, respectively.

2(d)(ii) The balance equations imply the followings:

$$p_1 = (\lambda_0/\mu_1)p_0$$

$$p_{n+1} = (\lambda_n/\mu_{n+1})p_n + (1/\mu_{n+1})(\mu_n p_n - \lambda_{n-1}p_{n-1})$$

By induction, $p_{n+1} = (\lambda_n/\mu_{n+1})p_n$.

Consequently, $p_n = c_n p_0$ with $c_n = (\lambda_0\lambda_1 \cdots \lambda_{n-1}) / (\mu_1\mu_2 \cdots \mu_n)$

Notice that $1 = p_0 + p_1 + p_2 + \ldots = p_0 + c_1 p_0 + c_2 p_0 + \ldots$ Hence $p_0 = 1/(1 + c_1 + c_2 + \ldots)$.

Therefore the steady-state probability distribution $p_0, p_1, \ldots$, is represented as functions of birth rates and death rates.

4(a)(i) Let X_i be the weekly observations/demand in the time series, Y_i the two-week moving average in period i. Then $Y_i = (X_{i-2} + X_{i-1})/2$. Therefore $Y_6 = (X_5 + X_4)/2 = 23.5$.

4(a)(ii) Let F_i be the forecast of period i by applying exponential smoothing. Then $F_i = \alpha X_{i-1} + (1-\alpha) F_{i-1}$, where $\alpha = 0.8$ is the smoothing parameter. Let $F_1 = X_1$ for initialisation. Then $F_1 = 13$, $F_2 = 13$, $F_3 = 16.2$, $F_4 = 18.44$, $F_5 = 22.088$, $F_6 = 23.6176$. Therefore the forecast for demand in week 6 is 23.6176.

4(a)(iii) The global forecasting fitting error should be used to measure effectiveness of a forecasting method. Some concrete criteria of measuring the fitting error are mean absolute deviation (MAD), mean squared error (MSE) and mean absolute percentage error (MAPE).

4(b)(i) From the summary output, the intercept is -16.225062 and the slope (asking price) is 0.000528. Therefore the regression equation reads

$$y = -16.225062 + 0.000528\,x,$$

where x and y represent the asking price and the predicted number of weeks on the market before the home is sold. The slope indicates that on average, the change of the number of weeks (in a fractional number) the home is on the market before it is sold for any unit change of price. For example, if the asking price for a home is increased by £1000, then on average the same home will be on the market with additional 0.528 weeks before it is sold. Here the intercept does not have an intuitive interpretation as it states that the number of weeks on the market is negative when the asking price is zero (no home is priced as zero).

4(b)(ii) The correlation coefficient for two random variables X and Y is defined by

$$\text{Covar}(X,Y)/(S_x\, S_y)$$

where Covar(X,Y) is the covariance between X and Y, and S_x and S_y are the standard deviation of X and Y respectively. The correlation coefficient of two random variables measures the strength of their linear relationship. It is a number between -1 and 1. The closer the correlation coefficient is to 1 or -1, the stronger linear relationship that the two random variables have. The closer the correlation coefficient is to 0, the weaker linear relationship that the two random variables have. From the summary table, we see that the correlation coefficient for this example is 0.705948.

4(b)(iii) A 95% confidence interval for the slope is [0.000237, 0.000820] from the summary output. A 95% confidence interval for the slope means that we are sure that the slope will be in the confidence interval with a 95% of the chance.

4(b)(iv) When a home price is x = £100,000, $y = -16.225062 + 0.000528\,x = 36.59$ weeks. A rough 95% confidence interval for y is $[y - 2\,S_e, y + 2\,S_e] = [16.66, 60.52]$ weeks, where S_e is the standard error in the summary table.

4(b)(v) It is not recommended to use this regression equation to predict the number of weeks on the market before it is sold for a home with a price about £250,000 as this price is much higher than the maximum price $150,000 in the sample. It is dangerous to do extrapolations using the regression equation.

4(b)(vi) The slope and the intercept parameters in the regression equation are derived from solving a least square optimisation problem which is formulated using the sample historical data. Clearly the optimal values for the slope and the intercept depend on the sample data. In other words, values for the slope and the intercept are expected to change when a different set of sample data is used. In general, the slope and the intercept are random variables, which are of student's t-distributions and approximately of Normal distributions when the sample size is large. In the following we derive the mathematical formulae for the expected mean and the standard deviation of the slope. Assume the true linear relationship between the dependent and independent variable is
$y = \alpha + \beta x + \varepsilon$
where ε represents the error term and is of a Normal distribution with mean 0 and standard deviation σ. Let $E(b)$ and Var(b) be the mean and standard deviation of the slope. Then we have

$$b = \sum \lambda_i(\alpha + \beta x_i + \varepsilon_i) = \alpha \sum \lambda_i + \beta \sum \lambda_i x_i + \sum \lambda_i \varepsilon_i$$

$$= \beta + \sum \lambda_i \varepsilon_i \quad (\text{where } \lambda_i = \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2})$$

Therefore,

$$E(b) = E(\beta + \sum \lambda_i \varepsilon_i) = \beta$$

$$\text{Var}(b) = \text{Var}(\sum \lambda_i \varepsilon_i) = \sum \text{Var}(\lambda_i \varepsilon_i)$$

$$= \sum \lambda_i^2 \text{Var}(\varepsilon_i) = \sigma^2 \sum \lambda_i^2$$

$$= \sigma^2 \sum \left(\frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right)^2 = \sigma^2 / \sum (x_i - \bar{x})^2$$

4(c) In the simple linear regression, we assume that $y = \alpha + \beta x + \varepsilon$ is the relationship between x and y. That is, y can be represented as the sum of a linear component of x and an error term ε. Here

- x is the independent/explanatory variable
- y is the dependent variable
- β is a constant representing the slope
- α is a constant representing the intercept
- ε is the error term

It is further assumed that

- Error term is normally distributed with mean 0 and standard deviation σ.
- The normal distributions of the error term have the same mean and the same standard deviation for all given values of x in consideration.
- Random variables of the error term are mutually statistically independent.