SECTION ANSWERS

1 (a) Utility theory is useful in case where we are not only interested in the expected payoff but also in the distribution of the yields. It allows a decision maker to incorporate his risk preferences into the decision making process.

The utility is a function $U(x)$ that ranks each possible yield in the payoff set. Given two random yields Y_1 and Y_2, a decision maker prefers Y_1 to Y_2, if and only if the expected utility of Y_1 is larger than the expected utility of Y_2.

$$Y_1 \succ Y_2 \iff E(U(Y_1)) > E(U(Y_2))$$

The form of the utility function $U(x)$ defines the risk preference of the decision maker.

(b) :

- Given yields y_i with associated probabilities p_i the Expected Monetary Value (EMV) is given by

$$EMV = \sum_i y_i p_i$$

- A risk neutral investment strategy is indifferent between payouts with the same EMV, irrespective of risk.
- The CARA - constant absolute risk aversion utility is of the form

$$U(x) = 1 - e^{-ax}$$

(or a linear transformation of this). a is the constant of absolute risk aversion.

(c) The insurance risk premium β is the amount that a risk-averse decision maker would be willing to pay in order to avoid a fair gamble with mean μ.

β is related to the utility and the variance of the yields by

$$\beta \approx \frac{-U''(\mu)}{2U'(\mu)} Var(Y)$$

(d)

$$U(E(Y)) = U(\mu) = -e^{-a\mu}$$

For $E(U(Y))$ we need to know the distribution. Since the distribution is normal, the density function is

$$p(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(\frac{-(y-\mu)^2}{2\sigma^2})$$

Thus

$$E(U(Y)) = \int_{-\infty}^{\infty} U(y)p(y)dy = \int_{-\infty}^{\infty} -\frac{1}{\sqrt{2\pi\sigma^2}} \exp(-a\mu - \frac{(y-\mu)^2}{2\sigma^2})dy$$

Using the Hint with $A = 1/\sqrt{(}2\sigma^2), B = -\mu/\sqrt{(}2\sigma^2), C = -a$, this expression becomes

$$E(U(Y)) = -\exp(-a\mu + \frac{a^2}{2\sigma^2})$$

(e) Since β satisfies

$$E(U(Y)) = U(\mu - \beta)$$

we find

$$-\exp(-a\mu + \frac{a^2}{2\sigma^2}) = -\exp(-a\mu + a\beta)$$

i.e.

$$\beta = \frac{a\sigma^2}{2}$$

(f) The approximation gives

$$\beta \approx \frac{1}{2} \frac{a^2 e^{-a\mu}}{ae^{-a\mu}} \sigma^2 = \frac{a\sigma^2}{2}$$

which is exactly what we found above.

2 (a) :

- Simple queuing systems are conventionally labelled by

$$U/V/s/\kappa/W$$

– U and V denote the inter-arrival and service time distributions.

– s is the number of servers

– κ is the system capacity.

– W is the queuing protocol or discipline.

κ and W are optional with default values $\kappa = \infty$ and $W = FIFO$.

- Imagine a $G/G/s$ queue with arrival rate λ and service rate μ. The utilization factor (or traffic intensity) ρ is defined by

$$\rho = \frac{\lambda}{s\mu}$$

The interpretation of this is that ρ is the fraction of time we expect the service facility to be busy.
If $\rho > 1$ then the queue explodes, i.e. the number of people in the queue tends to infinity as $t \to \infty$.

- Exponential arrival times are often used because it is reasonable to assume that arrival times are memoryless. The probability that a new customer will arrive in the next 10 minutes, say, is independent of whether there has been a customer in the 10 minutes before that or not.

- The state $N(t)$ of a queuing system at time t is the number of customers in the system (i.e in the queue or in service) at time t. The system is said to be in a steady state if $P(N(t) = n)$ does not change with t anymore.

- Little's formula says that the average number of customers L in any steady state system (over some time interval) is equal to their average arrival rate, λ, multiplied by their average time in the system W.

$$L = \lambda W$$

(b) (i) This is the queue length $L_q = \sum_{i=1}^{3} p(i+2)i = 0.85$.

(ii) Customers are turned away if the queue is full, i.e. $n = 5$. The probability of this is 0.1.

(iii) Use Little's formula. The expected number of callers in the system is $L = 2.5$. Since $\lambda = 4$, $W = 2.5/4 = 0.625$ hours.

(iv) $W = W_q + \frac{1}{\mu}$, thus $L = L_q + \frac{\lambda}{\mu}$. Hence

$$\frac{\lambda}{\mu} = L - L_q = 2.5 - 0.85 = 1.65.$$

and

$$\rho = \frac{\lambda}{2\mu} = 0.825$$

(v) This is equivalent to (1 - the percentage of time that they are not busy), i.e. (1- the utilization faction) = 1-0.825 = 0.175.

(vi) They are both off the phone if there are no customers in the system, i.e. $n = 0$. The probablility of this is 0.1.

3 (a) :

(i) NO - The car never drives back to where it came from, hence it is not memoryless.

(ii) NO - For the same reason.

(iii) YES - The car has no memory of where it came from.

(iv) NO - The car needs to remember the last junction it came from, hence it is not memoryless.

(v) YES - The probability changes depending on what is NORTH, which is a parameter of the system, not the past history of the car.

(b) :

(i) A matrix P is stochastic if every row is a distribution, i.e.

- $0 \leq p_{ij} \leq 1$
- $\sum_j p_{ij} = 1$.

M is stochastic.

(ii) Label the points by $\{a,b,c,d,e,f\}$.

There are two classes $\{a,b,c,e,f\}$ and $\{d\}$. The first one is absorbing/closed.

(iii) $\mathbf{q}M^2 = (3/8, 0, 1/4, 0, 3/8, 0)$

(iv) $\mathbf{u}M = \mathbf{u}$ gives us the following solution

$$\mathbf{u} \propto (9, 8, 12, 0, 6, 8)$$

Since $\mathbf{u}$ needs to be a distribution the u_i need to sum to 1. We can achieve this by dividing the above by 43

$$\mathbf{u} = \frac{1}{43}(9, 8, 12, 0, 6, 8)$$

u is the limiting distribution because the matrix M is not periodic.

(v) The expected return time for state d is ∞. The return time for any other state is $\frac{1}{u_i}$, thus the shortest return time is $\frac{43}{12}$ for state c.

4 (a) In order to find the parameters a and b that minimize the sum of squared errors (SSE) $\sum_i (a+bx_i-y_i)^2$we need to differentiate the SSE with respect to the parameters a and b. Setting $\frac{\partial S}{\partial a} = \frac{\partial S}{\partial b} = 0$ gives us the minimum of the SSE.

$$\frac{\partial S}{\partial a} = \sum_i 2(a+bx_i-y_i) = 0$$

gives

$$a = \bar{y}$$

and

$$\frac{\partial S}{\partial b} = \sum_i 2(a+bx_i-y_i)x_i = 0$$

gives

$$0 = \sum_i (\bar{y}+bx_i-y_i)x_i$$

$$\sum_i (y_i-\bar{y})x_i = b\sum_i x_i^2$$

Thus we find

$$b = \frac{\sum_i x_i y_i}{\sum_i x_i^2}$$

(b) The expected value $E(y_i) = \alpha + \beta E(x_i)$. Thus

$$E(a) = E(\bar{y}) = \alpha + \beta E(\bar{x}) = \alpha$$

and

$$\begin{aligned} E(b) &= \frac{\sum_{i=1}^n x_i(E(y_i)-E(\bar{y}))}{\sum_i x_i^2} \\ &= \frac{\sum_{i=1}^n x_i((\alpha+\beta x_i)-(\alpha+\beta\bar{x}))}{\sum_i x_i^2} \\ &= \beta\frac{\sum_{i=1}^n x_i(x_i-\bar{x})}{\sum_i x_i^2} \\ &= \beta \end{aligned} \qquad (1)$$

(c) The distributions for a and b are normal with mean α and β respectively. If we believe that our model is correct, then the observations y_i have been generated by $y_i = \alpha + \beta x_i + \varepsilon_i$ where all the ε_i are independently drawn from the same distribution.

Version: 1

We can therefore rewrite the expression for the slope as

$$b = \frac{\sum_i x_i(\alpha + \beta x_i + \varepsilon_i)}{\sum_i x_i^2} = \beta + \frac{\sum_i x_i \varepsilon_i}{\sum_i x_i^2}$$

Define $S_{XX} = \sum_i x_i^2$. According to the central limit theorem, when n is large enough, b has a normal distribution with mean β and variance σ^2/S_{XX} where σ^2 is the unknown variance of the ε_i for all i.

$$b \sim N\left(\beta, \frac{\sigma^2}{S_{XX}}\right)$$

The standard error

$$S_e = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2}}$$

is an unbiased estimator for the parameter σ.

(d)

$$\frac{(b-\beta)\sqrt{S_{XX}}}{S_e} \sim t_{n-2}$$

We can look up the values from a t-distribution to give us confidence intervals for β. A $100(1-\gamma)\%$ confidence interval can be constructed by

$$[b - t_{n-2}(\gamma/2)\sigma_b, b + t_{n-2}(\gamma/2)\sigma_b]$$

where $\sigma_b = \frac{S_e}{\sqrt{S_{XX}}}$ and $t_n(\gamma)$ is defined by $P(T > t_n(\gamma)) = \gamma$.

(e) A prediction interval gives a confidence interval that for a given value of x the y value lies in a certain range about the predicted value.

We have $y = a + bx + \varepsilon = \bar{y} + bx + \varepsilon$. Thus

$$\begin{aligned} Var(y) &= Var(\bar{y}) + x^2 Var(b) + Var(\varepsilon) \\ &= \frac{\sigma^2}{n} + x^2 \frac{\sigma^2}{S_{XX}} + \sigma^2 \\ &= \left(1 + \frac{1}{n} + \frac{x^2}{S_{XX}}\right)\sigma^2 \end{aligned}$$

and since S_e is an unbiased estimator for σ we find the standard deviation of y by

$$\sigma_y(x) = \sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{S_{XX}}} S_e$$

A $100(1-\gamma)\%$ prediction interval can thus be constructed by

$$[\hat{y} - t_{n-2}(\gamma/2)\sigma_y, \hat{y} + t_{n-2}(\gamma/2)\sigma_y]$$