**APP MATH 3020 Stochastic Decision Theory**
**Assignment 1**

**Due: Friday, 17 August, 2018, 4 p.m. (Week 4).** $\boxed{\text{Total marks: 38}}$

**Question 1** $\boxed{\text{2 marks}}$

Make sure that in all your answers you

$\boxed{1/2}$ (a) use full and complete sentences.

$\boxed{1/2}$ (b) include units where necessary.

$\boxed{1/2}$ (c) use logical arguments in your answers and proofs.

$\boxed{1/2}$ (d) structure your answers and assignment clearly and precisely.

**Question 2** $\boxed{\text{10 marks}}$

Consider the following linear program:

$$\begin{array}{lll}
\text{(P)} & \text{maximise} & z = -5x_1 + x_2 - 4x_3 \\
& \text{such that} & 2x_1 + 2x_2 - 4x_3 = 1 \\
& & 2x_1 + 2x_2 + 2x_3 = 4 \\
& & x_1,\ x_2,\ x_3 \geq 0.
\end{array}$$

$\boxed{4}$ (a) Construct its dual using the general relationship between primal and dual (instead of rewriting the above LP in standard form and then writing the dual of the LP in the standard form).

**Solution:** The dual for the above LP is

$$\begin{array}{lll}
\text{(D)} & \text{minimise} & w = y_1 + 4y_2 \ [1] \\
& \text{such that} & 2y_1 + 2y_2 \geq -5 \\
& & 2y_1 + 2y_2 \geq 1 \\
& & -4y_1 + 2y_2 \geq -4 \\
& & y_1, y_2 \text{ free. } [1]
\end{array}$$

[2 if all three inequalities are correct; 1 if at least one (and at most two) is wrong]

$\boxed{6}$ (b) Provide the optimal solutions of both the primal and dual, using `linprog.m`. Include your MATLAB code and the output of the code.

**Solution:** As `linprog` requires the objective function to be a minimisation, we first convert the objective function of the primal to minimise $-z = 5x_1 - x_2 + 4x_3$.

For the primal, one may use the following code: [1 for both code and output]

```
% primal objective
f = [5; -1; 4];
% equalities
Aeq = [2 2 -4;
       2 2 2];
```

```
beq = [1; 4];
% lower bounds for x_i
LB = [0; 0; 0];
% solving the primal
[X, fval, exitflag] = linprog(f,[],[],Aeq,beq,LB)
```

The output for the above code is

```
Optimization terminated.
X =
    0.0000
    1.5000
    0.5000
fval =
    0.5000
exitflag =
     1
```

As the exit flag = 1, this implies `linprog` has converged to a solution, $X$. The output indicates that $\boldsymbol{x}^* = (0, 1.5, 0.5)$ [1] , with optimal value $z^* = -0.5$. [1]
[0 mark for the objective function if $z = 0.5$]

For the dual, in order to use `linprog` we first have to convert the $\geq$ inequalities into $\leq$. One may use the following code: [1 for both code and output]

```
% dual objective
f = [1; 4];
% inequalities
A = [-2 -2;
     -2 -2;
      4 -2];
b = [5; -1; 4];
% solving the dual
[Y, fval, exitflag] = linprog(f,A,b,[],[])
```

The output for the above code is

```
Optimization terminated.
Y =
    0.8333
   -0.3333
fval =
   -0.5000
exitflag =
     1
```

The output indicates $\boldsymbol{y}^* = (0.833, -0.333)$ [1] , with optimal value $w^* = -0.5$. [1] So the optimal values of the primal and dual are equal, as one would expect from the strong duality theorem.

(Note: it was not necessary to include the first inequality constraint in the dual, $2y_1 + 2y_2 \geq -5$, as this is automatically satisfied if the second inequality holds.)

**Question 3** $\boxed{\text{16 marks}}$

Consider the following linear program:

$$
\text{(P)} \qquad \begin{array}{ll} \text{maximise} & z = 4x_1 + 8x_2 + 5x_3 \\ \text{such that} & x_1 \geq 10 \; (= b_1) \\ & x_2 \geq 9 \; (= b_2) \\ & x_3 \geq 3 \; (= b_3) \\ & 2x_1 + 3x_2 + x_3 \leq 80 \; (= b_4) \\ & x_1 + 2x_2 + 2x_3 \leq 70 \; (= b_5) \\ & x_1,\, x_2,\, x_3 \geq 0. \end{array}
$$

Suppose we replace $b_i$ with $b_i(\boldsymbol{\zeta})$ for $i = 1, 2, \ldots, 5$, with

$$
\boldsymbol{\zeta} = \begin{cases} \boldsymbol{\varepsilon}_1 & \text{with probability } 0.3, \\ \boldsymbol{\varepsilon}_2 & \text{with probability } 0.5, \\ \boldsymbol{\varepsilon}_3 & \text{with probability } 0.2, \end{cases}
$$

where

$$
\boldsymbol{b}(\boldsymbol{\zeta}) = \begin{pmatrix} b_1(\boldsymbol{\zeta}) \\ b_2(\boldsymbol{\zeta}) \\ b_3(\boldsymbol{\zeta}) \\ b_4(\boldsymbol{\zeta}) \\ b_5(\boldsymbol{\zeta}) \end{pmatrix}, \quad \text{with} \quad \boldsymbol{b}(\boldsymbol{\varepsilon}_1) = \begin{pmatrix} 8 \\ 6 \\ 1 \\ 80 \\ 70 \end{pmatrix}, \quad \boldsymbol{b}(\boldsymbol{\varepsilon}_2) = \begin{pmatrix} 10 \\ 10 \\ 3 \\ 80 \\ 70 \end{pmatrix} \quad \text{and} \quad \boldsymbol{b}(\boldsymbol{\varepsilon}_3) = \begin{pmatrix} 13 \\ 11 \\ 6 \\ 80 \\ 70 \end{pmatrix}.
$$

Assume that we can meet any shortfall in demand through recourse at market, but we must pay $q_1, q_2$ and $q_3$ (units of currency) per unit of $x_1, x_2$ and $x_3$ purchased at market, respectively.

$\boxed{6}$ (a) Formulate and write down for this problem, a two-stage stochastic linear program (SLP) with fixed recourse.

**Solution:** First, note that the values of $b_4(\boldsymbol{\zeta})$ and $b_5(\boldsymbol{\zeta})$ remain unchanged (80 and 70, respectively), regardless of the realisation of $\boldsymbol{\zeta}$. So the fourth and fifth constraints are in fact deterministic. Second, as the objective function is presently a maximisation and we would like to minimise the expected recourse, we need to *subtract* the expected recourse from the current objective function.

The two-stage stochastic LP with fixed recourse can be written as follows:

$$
\begin{array}{ll} \text{maximise} & 4x_1 + 8x_2 + 5x_3 - \mathbb{E}_{\boldsymbol{\zeta}}[Q(\boldsymbol{x}, \boldsymbol{\zeta})] \text{ [1]} \\ \text{such that} & 2x_1 + 3x_2 + x_3 \leq 80 \\ & x_1 + 2x_2 + 2x_3 \leq 70 \text{ [1 mark for both deterministic constraints]} \end{array}
$$

where $\mathbb{E}_{\boldsymbol{\zeta}}[Q(\boldsymbol{x}, \boldsymbol{\zeta})] = 0.3Q(\boldsymbol{x}, \boldsymbol{\varepsilon}_1) + 0.5Q(\boldsymbol{x}, \boldsymbol{\varepsilon}_2) + 0.2Q(\boldsymbol{x}, \boldsymbol{\varepsilon}_3)$, [1] and

$$
\begin{array}{rl} Q(\boldsymbol{x}, \boldsymbol{\varepsilon}_1) = & \min q_1y_1 + q_2y_2 + q_3y_3 \\ \text{such that} & x_1 + y_1 \geq 8 \\ & x_2 + y_2 \geq 6 \\ & x_3 + y_3 \geq 1 \\ & y_1, y_2, y_3 \geq 0; \end{array}
$$

$$
\begin{aligned}
Q(\boldsymbol{x}, \boldsymbol{\varepsilon}_2) = \min\, & q_1 y_4 + q_2 y_5 + q_3 y_6 \\
\text{such that} \quad & x_1 + y_4 \geq 10 \\
& x_2 + y_5 \geq 10 \\
& x_3 + y_6 \geq 3 \\
& y_4, y_5, y_6 \geq 0; \\
Q(\boldsymbol{x}, \boldsymbol{\varepsilon}_3) = \min\, & q_1 y_7 + q_2 y_8 + q_3 y_9 \\
\text{such that} \quad & x_1 + y_7 \geq 13 \\
& x_2 + y_8 \geq 11 \\
& x_3 + y_9 \geq 6 \\
& y_7, y_8, y_9 \geq 0.
\end{aligned}
$$

[3 marks, one for each of the above LPs]

---

Alternatively, one can also write the two-stage SLP as follows:

$$
\begin{aligned}
\text{maximise} \quad & 4x_1 + 8x_2 + 5x_3 - \mathbb{E}_{\boldsymbol{\zeta}}\left[q_1 y_1(\boldsymbol{\zeta}) + q_2 y_2(\boldsymbol{\zeta}) + q_3 y_3(\boldsymbol{\zeta})\right] \text{ [1]} \\
\text{such that} \quad & 2x_1 + 3x_2 + x_3 \leq 80 \\
& x_1 + 2x_2 + 2x_3 \leq 70 \text{ [1 mark for both deterministic constraints]} \\
& x_1 + y_1(\boldsymbol{\zeta}) \geq b_1(\boldsymbol{\zeta}) \text{ [1]} \\
& x_2 + y_2(\boldsymbol{\zeta}) \geq b_2(\boldsymbol{\zeta}) \text{ [1]} \\
& x_3 + y_3(\boldsymbol{\zeta}) \geq b_3(\boldsymbol{\zeta}) \text{ [1]} \\
& y_1(\boldsymbol{\zeta}), y_2(\boldsymbol{\zeta}), y_3(\boldsymbol{\zeta}) \geq 0. \text{ [1]}
\end{aligned}
$$

(In that case, all of the constraints must be satisfied with probability 1.)

$\boxed{4}$ (b) Write the above SLP in the extended form.

**Solution:**

$$
\begin{aligned}
\text{maximise} \quad & 4x_1 + 8x_2 + 5x_3 - 0.3(q_1 y_1 + q_2 y_2 + q_3 y_3) \\
& - 0.5(q_1 y_4 + q_2 y_5 + q_3 y_6) - 0.2(q_1 y_7 + q_2 y_8 + q_3 y_9) \text{ [1]} \\
\text{such that} \quad & 2x_1 + 3x_2 + x_3 \leq 80 \\
& x_1 + 2x_2 + 2x_3 \leq 70 \text{ [1 mark for the first two constraints]} \\
& x_1 + y_1 \geq 8 \\
& x_2 + y_2 \geq 6 \\
& x_3 + y_3 \geq 1 \\
& x_1 + y_4 \geq 10 \\
& x_2 + y_5 \geq 10 \\
& x_3 + y_6 \geq 3 \\
& x_1 + y_7 \geq 13 \\
& x_2 + y_8 \geq 11 \\
& x_3 + y_9 \geq 6 \text{ [1 mark for the next nine constraints]} \\
& y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8, y_9 \geq 0. \text{ [1]}
\end{aligned}
$$

[6] (c) Solve the recourse DEP in part (b) for $q_1 = 10, q_2 = 1$ and $q_3 = 20$. Provide MATLAB code and output of the code, with an interpretation.

**Solution:** We can solve the recourse DEP using the following MATLAB code: [2]

```
q_1 = 10; q_2 = 1; q_3 = 20;
% objective function
f = -[4; 8; 5; -0.3*q_1; -0.3*q_2; -0.3*q_3; -0.5*q_1;-0.5*q_2; -0.5*q_3;...
           -0.2*q_1;-0.2*q_2; -0.2*q_3];
% inequalities
A = [ 2  3  1  0  0  0  0  0  0  0  0  0; % deterministic constraint #1
      1  2  2  0  0  0  0  0  0  0  0  0; % deterministic constraint #2
     -1  0  0 -1  0  0  0  0  0  0  0  0; % -x_1 - y_1
      0 -1  0  0 -1  0  0  0  0  0  0  0; % -x_2 - y_2
      0  0 -1  0  0 -1  0  0  0  0  0  0; % -x_3 - y_3
     -1  0  0  0  0  0 -1  0  0  0  0  0; % -x_1 - y_4
      0 -1  0  0  0  0  0 -1  0  0  0  0; % -x_2 - y_5
      0  0 -1  0  0  0  0  0 -1  0  0  0; % -x_3 - y_6
     -1  0  0  0  0  0  0  0  0 -1  0  0; % -x_1 - y_7
      0 -1  0  0  0  0  0  0  0  0 -1  0; % -x_2 - y_8
      0  0 -1  0  0  0  0  0  0  0  0 -1];% -x_3 - y_9
b = [80; 70; -8; -6; -1; -10; -10; -3; -13; -11; -6];
% solving the recourse DEP
[X, fval, exitflag] = linprog(f,A,b,[],[],[zeros(12,1)])
```

The output of the above code is: [1]

```
Optimization terminated.
X =
   13.0000
   12.7500
   15.7500
    0.0000
    0.0000
    0.0000
    0.0000
    0.0000
    0.0000
    0.0000
    0.0000
    0.0000
fval =
 -232.7500
exitflag =
     1
```

This indicates that the optimal solution is $\boldsymbol{x}^* = (13, 12.75, 15.75)$, with the optimal value of 232.75. [1 mark for both optimal solution and value] . No recourse was needed. This is logical, as (a) the penalties for buying additional units are high, (b) there are no penalties for overproducing, (c) we want to maximise the profits from $x_i$. So it is optimal to produce as much as possible to meet the worse-case scenario in terms of $\boldsymbol{b}$ (while satisfying deterministic constraints). [2]

**Question 4** $\boxed{\text{10 marks}}$

Suppose $X$ is a continuous random variable with density function

$$f(x) = \lambda^2 x e^{-\lambda x} \quad \text{for } x \geq 0.$$

Suppose $\lambda = 3$.

$\boxed{2}$ (a) Find the 99% CI for $X$.

**Solution:** We would like to determine an upper bound $b$ such that

$$P(X \in [0, b]) = 0.99.$$

This implies

$$\int_0^b \lambda^2 x e^{-\lambda x} dx = 0.99,$$

Let $u = x$ and $dv = -\lambda e^{-\lambda x}$. Then, $du = dx$, $v = e^{-\lambda x}$, and integrating by parts gives

$$\begin{aligned}\int_0^b \lambda^2 x e^{-\lambda x} dx &= -\lambda \left( x e^{-\lambda x}\Big|_0^b - \int_0^b e^{-\lambda x} dx \right) \\ &= -\lambda \left( b e^{-\lambda b} - \frac{e^{-\lambda x}}{-\lambda}\Big|_0^b \right) \\ &= -\lambda b e^{-\lambda b} - e^{-\lambda b} + 1 \\ &= 1 - e^{-\lambda b}(1 + \lambda b).\end{aligned}$$

Thus, we want to solve

$$1 - e^{-\lambda b}(1 + \lambda b) = 0.99. \text{ [1]}$$

This is not an easy equation to solve analytically, so we can solve it numerically using, for example, either the bisection method (taking advantage of the fact that the distribution function $F(x)$ is increasing) or simply plotting the graph on MATLAB and then using the data cursor. For $\lambda = 3$, $b \approx 2.2128$. [1]

$\boxed{5}$ (b) Determine a discrete approximation for $X$, with 10 realisations. Explain clearly how you obtain the realisations and their respective probabilities. (Include computer code if you have used a computer to help generate your realisations.)

**Solution:** We can follow the method outlined in the lecture notes:

1. Subdivide the interval $[0, 2.2128]$ into ten subintervals with equal lengths (each is of length 0.22128).

2. Sample, say, 100,000 realisations of $X$ over the interval $[0, 2.2128]$.

3. For each subinterval $i$, $i = 1, \ldots, 10$, take the average of the samples that fall within this subinterval. This gives us the $i$th realisation of the discrete approximation.

4. For $i = 1, \dots, 10$, find the relative frequency for the realisation $i$ by dividing the number of sample points within the $i$th subinterval by $100,000$.

If we use this method, the hardest part is Step 2, generating samples of $X$. We could do using the inverse method: generating $U(0,1)$ and finding $x$ such that $F(x) = U$ (using the same numerical procedure we used above to find $F(b) = 0.99$).

Alternatively, an easier way to generate samples of $X$ is as follows. Note that $X$ is an Erlang$(2,3)$, that is, the sum of two exponential random variables each with rate $\lambda = 3$. So we can also generate two samples of an Exp(3), which we can do easily using the inverse method, and its sum gives us a sample of $X$.

[2 marks for realisations, 1 marks for weights, 2 marks for explanations]

Of course one can also use many other methods to generate the samples as well as their weights. For example, each sample could be taken to be the mean from each subinterval (i.e. the mean of $X$ given that $X$ belongs to that subinterval), and the corresponding weight the probability mass of that subinterval.

2 (c) Propose a mathematical formula/method for assessing the error of your discretisation.

**Solution:** One error measurement is based on the difference between the cumulative distribution function $F(x)$ of $X$, and the cumulative distribution function $\widetilde{F}_n(x)$ of $\widetilde{X}$, the discrete approximation:

$$\sup_{x \geq 0} |F(x) - \widetilde{F}_n(x)|,$$

where $n$ denotes the number of realisations that we use (in this case, $n = 10$). (This means we have applied the uniform norm on the difference between the two cumulative distributions.)

We could also use a mean-squared error:

$$\sum_{i=1,\dots n} p_i (F(x_i) - \widetilde{F}_n(x_i))^2,$$

where $p_i$ is the relative frequency of realisation $i$.

Yet another error measurement is to integrate over the area of differences between the two cumulative distributions:

$$\int_0^\infty |F(x) - \widetilde{F}_n(x)| dx.$$

[2 marks for any sensible suggestion]

1 (d) Given the above, suggest how you would reduce the approximation error.

**Solution:** We could increase both the numbers of intervals and of realisations. As these go to infinity, all of the above three errors tend to 0 with probability 1.

[1 mark for any sensible suggestion]