Institute of Actuaries of India

Subject CT4 – Models

March 2018 Examination

INDICATIVE SOLUTION

Introduction

The indicative solution has been written by the Examiners with the aim of helping candidates. The solutions given are only indicative. It is realized that there could be other points as valid answers and examiner have given credit for any alternative approach or interpretation which they consider to be reasonable.

Solution 1:

i), ii)

	State Space	Time Set
Counting Process	Discrete	Discrete or Continuous
General Random Walk	Discrete or Continuous	Discrete
Poisson Process	Discrete	Continuous
Markov Jump Chain	Discrete	Discrete
Markov Jump Process	Discrete	Continuous

[2.5+2.5]

iii)

a) Number of times the account has been overdrawn since it was opened – Poisson Process / Counting Process
b) Status (overdrawn, in credit) of the account on the last day of each month – Markov Jump Chain/ Counting Process
c) number of direct debits paid since the account was opened – Poisson
d) Status (overdrawn, in credit) of the account at any time since the account was opened. Markov Jump Process [2]

[7 marks]

Solution 2:

i)

a) Right censoring refers to a life ceasing to be observed prior to the event of interest occurring.

[1]

b) Type I censoring occurs when the censoring times are known in advance and lives will be considered censored on a pre-determined date regardless of whether the event of interest has occurred.

[1]

c) Random censoring refers to the time of censoring being a random variable such that censoring may occur as a random event prior to the event of interest.

[1]

ii)

a) It is right censoring as it removes information about whether the member subsequently may have died. It is NOT Type I censoring because time of leaving service is not known in advance. It is Random censoring for the same reason.

[1]

b) Right censoring occurs because the censoring means no information is available about whether the member would subsequently have lapsed. This is Type I censoring as it the date of retirement would be known in advance. Member reaching its retirement date is not a random variable.

[1]

c) It is right censoring as it removes information about whether the members subsequently may have died. It is NOT Type I censoring as non-renewability of the policy may not be known in advance. It is Random censoring for the same reason.

[1]

[6 marks]

Solution 3:

i)

Model fitting: this occurs after the family of model has been decided and concerns the estimation of the values of parameters. The set of parameters to be estimated is determined by the choice of model family.

Model verification: once the model has been fitted we need to check that the fitted process resembles what has been observed. Generally we produce simulations of the process, using the estimated parameter values, and compare them with the observations.

[2]

ii) The parameters required to be estimated to model this as Markov processes are:

- Rate of leaving state *i*, λ_i for each *i,*
- the jump chain transition probabilities, r_{ij} for $j \neq i$, where r_{ij} is the conditional probability that the next transition takes the chain to state *j* given that it is now in state *i*.

Assumptions of the Markov model for determination of these parameters:

- Duration of holding time in state *i* has exponential distribution with parameter determined only by *i* and is independent of anything that happened before the current arrival in state *i*,
- Destination of the next jump after leaving state *i* is independent of the holding time in state *i* and of anything that happened before the chain arrived in state *i*.

[4]

iii) The average duration of stay in any particular state i is given by $\frac{1}{\lambda_i}$

State	$1/\lambda_i$
1	6
2	40
3	30

The transition probabilities are as follows:

$r_{12} = \frac{3}{8}$, $r_{13} = \frac{5}{8}$, $r_{21} = \frac{1}{4}$, $r_{23} = \frac{3}{4}$, $r_{31} = \frac{7}{8}$ and $r_{32} = \frac{1}{8}$

Thus the generator matrix is

$$\begin{pmatrix} -1/6 & 3/48 & 5/48 \\ 1/160 & -1/40 & 3/160 \\ 7/240 & 1/240 & -1/30 \end{pmatrix}$$

[4]

[10 marks]

Solution 4:

i) h(t,zi) = $h_o(\mathrm{t})\exp(\beta_1 \mathrm{z1} + \beta_2 \mathrm{z2} + \beta_3 \mathrm{z3})$, where
h(t, zi) is the hazard at time t;
h0(t) is the baseline hazard;
β1 … β3 are regression parameters;
z_1 is a covariate which takes the value 1 if the client is Female, 0 otherwise;
z_2 is a covariate which takes the value 1 if Non-smoker, 0 otherwise;
z_3 is a covariate which takes the value 1 if the area of residence is Rural, 0 otherwise;

[3]

ii) The baseline hazard refers to the male smokers who belong to urban area.

[1]

iii) h(t,zi) = $h_0(t)\exp(\beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \beta_4 z_4)$, where β_4 is the regression parameter and z4 is the marital status covariate which takes the value 0 if unmarried otherwise 1.
The probability of lapse on 5th policy anniversary of married female is higher by 20%.
Hazard of lapsation of married woman is $h_0(5)\exp(0.065 - 0.035 + 0.012 + \beta_4)$ i.e $h_0(5)\exp(0.042 + \beta_4)$
Hazard of lapsation of unmarried woman is $h_0(5)\exp(0.042)$
Therefore 1.2 = $\frac{\mathrm{h0(5)\ exp\ (0.042+}\ \beta_4)}{\mathrm{h0(5)\ exp\ (0.042)}}$

β_4 = ln 1.2

β_4 = 0.1823

For a male, non-smoker from the city area the probability that he does not lapse the policy in first 5 year is 0.5, the sum of the parameters is 0-0.035+0+0= - 0.035 and the hazard is $h_0(5)\exp(-0.035)$.
So the probability that the contract is still in force is
0.5 = $\exp\{-\int_0^5 h_o(t)\exp(-0.035)\}dt$

So value of the Integration $\int_0^5 \mathrm{h_o(t)dt} = \frac{-\mathrm{In}0.5}{0.9656} = 0.7178$

For the married female, non-smoker from rural area, the sum of the parameters is 0.065-0.035+0.012+0.1823 = 0.2243
The probability that she does not lapse the policy for at least 5 years is

$\exp\{-\int_0^5 \mathrm{h_o(t)exp(0.2243)dt}\}$

= $\exp\{-\int_0^5 \frac{\mathrm{In}0.5}{0.9656} * \exp(0.2243)\mathrm{dt}\}$ replacing integrating factor from (1)

= $\exp\{1.25 * -0.7178\}$

= 0.4072 i.e. 40.72%

Alternative solution:

In case covariate of additional parameter is interchangeably used i.e. for unmarried 1 and for married 0, then the solution will differ as follows:
For married female, non-smoker from rural the sum of parameters is 0.065-0.035+0.012 = 0.042

The probability of not lapsing the policy is
= $\exp\{-\int_0^5 h_o(t)\exp(0.042)dt\}$
= $\exp\{1.0429 * -.7178\}$ replacing integrating factor from (1)
=$\exp\{-0.748\}$
=0.473 i.e 47.3%

[5]
[9 marks]

Solution 5:

i) Kaplan Meir estimate of S(t)=$\Pi(1-\lambda J)$

For males

tj	dj	cj	nj	$\frac{dj}{nj}$	$(1-\lambda j)$	S(t)= $\Pi(1-\lambda J)$
3	2	0	17	0.12	0.88	0.88
4	2	2	15	0.13	0.87	0.76
8	1	0	11	0.09	0.91	0.70
10	1	0	10	0.10	0.90	0.63
12	1	2	9	0.11	0.89	0.56
20	2	0	6	0.33	0.67	0.37
21	1	0	4	0.25	0.75	0.28

t	S(t)
0<=t<3	1
3<=t<4	0.88
4<=t<8	0.76
8<=t<10	0.70
10<=t<12	0.63
12<=t<20	0.56
20<=t<21	0.37
21<=t<23	0.28

For Females

tj	dj	cj	nj	$\frac{dj}{nj}$	$(1-\lambda j)$	S(t)= $\Pi(1-\lambda J)$
3	1	0	18	0.06	0.94	0.94
4	1	1	17	0.06	0.94	0.88
6	1	0	15	0.07	0.93	0.82
7	2	0	13	0.15	0.85	0.70
8	2	1	11	0.18	0.82	0.57
12	2	0	8	0.25	0.75	0.43
15	1	2	6	0.17	0.83	0.36
19	1	0	3	0.33	0.67	0.24

t	S(t)
0<=t<3	1
3<=t<4	0.94
4<=t<6	0.88
6<=t<7	0.82
7<=t<8	0.70
8<=t<12	0.57
12<=t<15	0.43
15<=t<19	0.36
19<=t<23	0.24

[5]

ii) Greenwood's formula can be used to determine variance

$$\text{Var(F(t))}= S(t)^2 \sum_{tj\le t}\left(\frac{dj}{nj*(nj-dj)}\right)$$

For males

t_j	d_j	n_j	S(t)	$\left(\frac{\text{dj}}{\text{nj} * (\text{nj} - \text{dj})}\right)$
3	2	17	0.8824	0.008
4	2	15	0.7647	0.01
8	1	11	0.6952	0.009
10	1	10	0.6257	0.011
12	1	9	0.5561	0.014
20	2	6	0.3708	0.083
21	1	4	0.2781	0.083

Variance (duration 4) = $0.7647^2 * (0.008 + 0.01)$ as $\text{S(t)} = 1 - \text{F(t)}$
=0.0105

Probability of death at duration of 4 months is 1-0.76 i.e 0.235

Probability of death with 95% CI = $(0.235 - \sqrt{0.0105} * 1.96, 0.235 + \sqrt{0.0105} * 1.96)$
$= (0.03, 0.44)$

For Females

t_j	d_j	n_j	S(t)	$\left(\frac{\text{dj}}{\text{nj} * (\text{nj} - \text{dj})}\right)$
3	1	18	0.9400	0.003

4	1	17	0.8836	0.004
6	1	15	0.8217	0.005
7	2	13	0.6985	0.014
8	2	11	0.5728	0.02
12	2	8	0.4296	0.042
15	1	6	0.3565	0.033
19	1	3	0.2389	0.167

Variance (duration 4) $= 0.8836^2 * (0.003 + 0.004)$
$=0.0054$
Probability of death at duration 4 months =1-0.8836=0.1164

Probability of death with 95% CI = $(0.1164 - \sqrt{0.0054} * 1.96, 0.1164 + \sqrt{0.0054} * 1.96)$
$= (-0.03, 0.26)$

[5]

iii) The probability of death of males with 95% CI ranges from 3% to 44% whereas the probability of death of females ranges upto 26% only, clearly male pensioners has higher mortality rates.

[2]
[12 marks]

Solution 6:

i) (Let $P'_x(t)$ be the number of policies inforce aged x nearest birthday at time t .
Also, let $P_x(t)$ be the number of policies inforce aged x last birthday at time t
Let E_x^C refers to the central exposed to risk at age label x respectively.
$E_x^C = \int_{t=0}^{2} P'x(t)dt$
Assuming that $P'_{56}(t)$ is linear over the year (2015,2016) and (2016,2017), we can approximate the exposure as follows
E_{56}^c = ½*(P'_{56}(2015)+ P'_{56}(2016))+ ½*(P'_{56}(2016)+ P'_{56}(2017))
= ½*P'_{56}(2015)+ P'_{56}(2016)+½*P'_{56}(2017)
Since, the number of policyholders aged label 56 nearest birthday will be between 55.5 and 56.5 i.e. between age label 55 last birthday and 56 last birthday. Assuming that the birthdays are uniformly distributed over the calendar year:

P'_{56}(2015) = ½*(P_{55}(2015)+ P_{56}(2015))
= 20050

Similarly,
P'_{56}(2016) = ½*(P_{55}(2016)+ P_{56}(2016))
= 20800

And,

$P'_{56}(2017) = ½*(P_{55}(2017)+ P_{56}(2017))$
$= 19250$

$E_{56}^{c} = ½*20050+20800+1/2*19250$
$= 40450$

$\mu_{56} = d_{56}/ E_{56}^{c}$
$= 1380/40450$
$= 0.0341$

Deriving the force of mortality for age 57 as above:

$P'_{57}(2015) = ½*(P_{56}(2015)+ P_{57}(2015))$
$= 19850$

Similarly,

$P'_{57}(2016) = ½*(P_{56}(2016)+ P_{57}(2016))$
$= 20900$

And,

$P'_{57}(2017) = ½*(P_{56}(2017)+ P_{57}(2017))$
$= 17500$

$E_{57}^{c} = ½*19850+20900+1/2*17500$
$= 39575$

$\mu_{57} = d_{57}/ E_{57}^{c}$
$= 1420/39575$
$= 0.03588$

dx is deaths aged x nearest birthday on the date of death. So the age label at death changes with reference to life year. Therefore the age at the middle of life year is x and estimates μ_{x}.

[6]

ii) We can estimate the initial rates of mortality using the estimated values of μ from part (i) and the following formula

$q_{55.5} = 1- \exp(-\mu_{56})$
$= 0.0335$

And

$q_{56.5} = 1- \exp(-\mu_{57})$
$= 0.0352$

[2]

[8 marks]

Solution 7:

i) Past history is needed to decide where to go in the chain.
If a sportsmen is at A and his/her performance reduces, you need to know what level of performance he was at the previous year to determine whether he or she drops one or two levels. [2]

ii) The B level needs to be split into two.

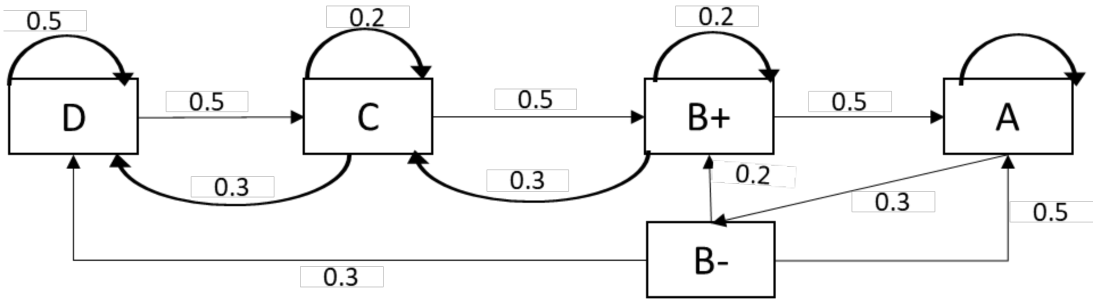
B+ is the level with no reduction in performance parameter last year

B- is the level with reduction in performance parameter last year

These levels are for modelling purposes and not the actual levels.

[2]

iii) The Process diagram is given below.



[2]

iv) The transition matrix is given by:

$$\begin{array}{c|ccccc|} & D & C & B+ & B- & A \\ \hline D & 0.5 & 0.5 & 0 & 0 & 0 \\ C & 0.3 & 0.2 & 0.5 & 0 & 0 \\ B+ & 0 & 0.3 & 0.2 & 0 & 0.5 \\ B- & 0.3 & 0 & 0.2 & 0 & 0.5 \\ A & 0 & 0 & 0 & 0.3 & 0.7 \end{array}$$

[2]

v) Stationary distribution:

The stationary disctribution is the set of probabilities that satisfy the martrix equation

π = πP with and additional condition $\sum \pi_i$ = 1

Written out fully, this set of matrix equation is

$0.5\pi_D + 0.3\pi_C + 0.3\pi_{B-} = \pi_D$ ---a)

$0.5\pi_D + 0.2\pi_C + 0.3\pi_{B+} = \pi_C$ ---b)

$0.5\pi_C + 0.2\pi_{B+} + 0.2\pi_{B-} = \pi_{B+}$ ---c)

$0.3\pi_A = \pi_{B-}$ ---d)

$0.5\pi_{B+} + 0.5\pi_{B-} + 0.7\pi_A = \pi_A$ ---e)

Substituting (d) in (e)

$0.5\pi_{B+} + 0.5 * 0.3\pi_A + 0.7\pi_A = \pi_A$

=> $\pi_{B+} = 0.3\pi_A$ …….(vi)

Substituing π_{B+} *and* π_A in (c) we get

$0.5\pi_C + 0.2 * 0.3\pi_A + 0.2 * 0.3\pi_A = 0.3\pi_A$

=> $\pi_C = 0.36\pi_A$

Substituing π_{B+}, π_C *and* π_A in (b)

$0.5\pi_D + 0.2 * 0.36\pi_A + 0.3 * 0.3\pi_A = 0.36\pi_A$ ---b)

=> $\pi_D = 0.396\pi_A$

Substiting all in $\sum \pi_i$ = 1

$\pi_A + 0.3\pi_A + 0.3\pi_A$+$0.36\pi_A$+$0.396\pi_A$=1

- ⇨ π_A = 0.42445
- ⇨ π_{B-} = 0.12733
- ⇨ π_{B+} = 0.12733
- ⇨ π_C = 0.15280
- ⇨ π_D = 0.1680 8

[3]

vi) Long run average contract value is
100% π_D + 120% π_C + 150% (π_{B+} + π_{B-}) + 175% π_A
= 1.4762 million USD

[2]

vii) Let m_i be the number of transitions (and years) taken to reach level A from any current level i. Then

$m_D = 1 + 0.5m_D + 0.5m_C$ .....(a)

$m_c = 1 + 0.2m_C + 0.5m_{B+} + 0.3m_D$ .....(b)

$m_{B+} = 1 + 0.2m_{B+} + 0.5m_A + 0.3m_C$ .....(c)

$m_{B-} = 1 + 0.3m_D + 0.5m_A + 0.2m_{B+}$ .....(c)

m_A = 0

from (a) we get,

$m_D = 2 + m_C$

substituting in (b) we get,

$m_c = 1 + 0.2m_C + 0.5m_{B+} + 0.3(2 + m_C)$

$m_{B+} = m_C - 3.2$

substituting in (c) ,

$m_{B+} = 1 + 0.2m_{B+} + 0.5m_A + 0.3m_C$

$m_C - 3.2 = 1 + 0.2(m_C - 3.2) + 0.3m_C$

m_C = 7.12 years

Therefore, m_D = 2+ m_C = 9.12 years

(note: this is an expected value and need not be rounded to integer number)

[4]

[17 marks]

Solution 8:

i)

The crude rates are subject to random sampling error. There may be irregularities in the progression of rate from one age to another. Graduation uses information from adjacent ages to improves the estimates at each age by reducing random sampling error and produce smooth set of rates.

Example: Mortality rates are used to compute financial quantities such as premiums for life insurance contracts. It is very desirable that such quantities progress smoothly with age. Irregular premiums at some ages may lead to anti- selection risk for example, more policies will be sold at ages where premium rates are the lowest (possibly underestimated). Hence the need for graduation. [2]

ii) Since there is no other decrement, the no. of pensioners at age x =no. of pensioners at age x-1 minus the no. of deaths between age x-1 and x.

The following table is produced:

Age	No. of deaths(A)	Ex	qx (as per pricing assumption)	Expected deaths (E)	A-E	(A-E)^2/E
60	14	10000	0.00128	12.80	1.20	0.11
61	16	9986	0.00144	14.38	1.62	0.18
62	13	9970	0.00162	16.15	-3.15	0.61
63	15	9957	0.00183	18.22	-3.22	0.57
64	17	9942	0.00206	20.48	-3.48	0.59
65	14	9925	0.00233	23.13	-9.13	3.60
66	30	9911	0.00263	26.07	3.93	0.59
67	25	9881	0.00297	29.35	-4.35	0.64
68	27	9856	0.00335	33.02	-6.02	1.10
69	42	9829	0.00378	37.15	4.85	0.63
70	35	9787	0.00425	41.59	-6.59	1.05
Total	248					9.68

Null Hypothesis: The Pricing table reflects the observed experience.

With 11 degrees of freedom, the value of tabular X^2 is 19.68 at 5% level which is greater than the calculated value X^2 which is 9.68.

So the result of the test would support the null hypothesis.

[5]

iii) Sign test
The null hypothesis is that the pricing assumption rates are the same observed rates
We have 7 negative signs out of 11 ages

P(7 Negative signs) = $\binom{11}{7}$* 0.5^11
=11!/(7!*4!)*0.5^11
=0.1611 i.e. 16.11%
Or
The probability of getting 7 or fewer negative signs is
1-(P(8)+p(9)+p(10)+p(11))
=1-0.1133
=0.8867 i.e. 88.67%
Which is greater than 0.025 (two tailed test)
Therefore at the 95% significance level we do not reject the hypothesis that the observed rates are consistent with pricing assumption.

[3]

iv)
Let N be the total no. of policyholders between age x and x+1.
Let π_i is the proportion of policyholders having i (=1,2,3 …) no. of policies
We have total no. of policies = $\sum_i i\pi_i N$, where i=1,2,3…

Let C, is the no. of claims
In case there are no dependencies, C could have been assumed to follow binomial distribution with parameter $(\sum_i i\pi_i N\, , q_x)$

In which case the Var(C) would have been $\sum_i i\pi_i N q_x\,(1-q_x)$ …Equation 1
Let Di is the no. of deaths among the πi*N lives each with i policies, and let Ci be the no. of claims among the same lives.
We can say that Di is Binomial ($\pi_i N$, q_x) as deaths are independent.

E[C]= $E[\sum_i Ci]$= $E[\sum_i iDi]$= $\sum_i iE[Di]$ = $\sum_i i\pi_i N q_x$

Var[C]= $Var[\sum_i Ci]$= $Var[\sum_i iDi]$= $\sum_i i^2 Var[Di]$

=$\sum_i i^2\pi_i N q_x\,(1-q_x)$ ..Equation 2

Ratio of equation 2 to equation 1

$$= \frac{\sum_i i^2\pi_i N q_x(1-q_x)}{\sum_i i\pi_i N q_x}$$

$$= \frac{\sum_i i^2\pi_i}{\sum_i i\pi_i}$$

[5]
[15 marks]

Solution 9:

i)

A Markov jump process is a continuous-time Markov process with a discrete state space. For a process to be Markov, the future development of the process must depend only on its current state. This is the case here, as the future of the process depends only on the number of consensus for the current transaction. The number of consensus for the current transaction also has a discrete state space {0, 1, 2, 3}.

(Note: that immediately after the 4th consensus, the next transaction from the queue is shared with the blockchain network nodes, so fourth consensus is not required for modelling purpose and all nodes must now work on new transaction to solve its cryptographic problem.)

[2]

ii) The generator matrix *A* is

$$\begin{vmatrix} -\beta & \beta & 0 & 0 \\ 0 & -\beta & \beta & 0 \\ 0 & 0 & -\beta & \beta \\ \beta & 0 & 0 & -\beta \end{vmatrix}$$

[1]

iii) Kolmogorov's forward equation can be written in compact form as

$$\frac{d}{dt}P(t) = P(t)A$$

Which are , for j =0

$$\frac{d}{dt}p_{i0}(t) = \beta p_{i3}(t) - \beta p_{i0}(t)$$

And for j = 1,2,3

$$\frac{d}{dt}p_{ij}(t) = \beta p_{i,j-1}(t) - \beta p_{ij}(t)$$

[3]

iv) Since the waiting times under a Poisson process are exponential, the expected waiting time between the arrival of consensus for the current transaction is 1/β minutes. Successive waiting times are independent, therefore the expected waiting time for a node for next transaction is

$$E(t) = \sum_{i=0}^{3} p_i \frac{3-i}{\beta},$$

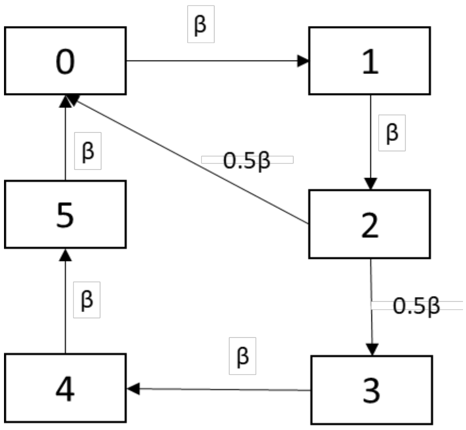
where $p_i$is the probability that the transaction has already received "i" consensus when the

new consensus arrives.
Since the $p_i s$ are all equal for *i* = 0, 1, 2, 3

$$E(t) = 0.25\left(\frac{3}{\beta}+\frac{2}{\beta}+\frac{1}{\beta}+\frac{0}{\beta}\right) = \frac{3}{2\beta} \text{ minutes}$$

[4]

v) The diagram and the transition matrix, *P*, is



$$\begin{vmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{vmatrix}$$

[2]

vi) The expected waiting time if the current transaction is less than 100k is:

$$E(t|3 - less\ than\ 100k\ Transaction) = \sum_{i=0}^{2} p_i \frac{2-i}{\beta}$$

$$= \frac{1}{3}\left(\frac{2}{\beta}+\frac{1}{\beta}+\frac{0}{\beta}\right) = \frac{1}{\beta}$$

The expected waiting time if the current transaction is more than 100k is:

$$E(t|6 - more\ than\ 100k\ Transaction) = \sum_{i=0}^{5} p_i \frac{5-i}{\beta}$$

$= \frac{1}{6}\left(\frac{5}{\beta} + \frac{4}{\beta} + \frac{3}{\beta} + \frac{2}{\beta} + \frac{1}{\beta} + \frac{0}{\beta}\right) = \frac{5}{2\beta}$

But 6-consensus transaction must expect to wait 2 times as long for addition to blockchain than 3-consensus transaction takes.

So when a consensus arrives for the current transaction, 2/3 (6/9) of the time the transaction at the front of the queue will be a transaction with value greater than 100K and only of the 1/3(3/9) time will is be a transaction with value less than 100k.

So the overall expected waiting time in minutes is

1/3*(E(t| 3 consensus transaction)) + 2/3*(E(t| 6 consensus transaction))

$= \frac{1}{3} * \frac{1}{\beta} + \frac{2}{3} * \frac{5}{2\beta} = \frac{2}{\beta}$

As this is longer than 3/2β, the time to add transaction to the blockchain has increased.

[4]

[16 marks]

****