INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

21st November 2012

Subject CT4 – Models

Time allowed: Three Hours (10.00 – 13.00 Hrs)

Total Marks: 100

INSTRUCTIONS TO THE CANDIDATES

1. ***Please read the instructions on the front page of answer booklet and instructions to examinees sent along with hall ticket carefully and follow without exception***

2. ***Mark allocations are shown in brackets.***

3. ***Attempt all questions, beginning your answer to each question on a separate sheet. However, answers to objective type questions could be written on the same sheet.***

4. ***Please check if you have received complete Question Paper and no page is missing. If so, kindly get new set of Question Paper from the Invigilator.***

AT THE END OF THE EXAMINATION
Please return your answer book and this question paper to the supervisor separately.

Q. 1) [i] Define the hazard function as proposed by Cox Regression Model briefly explaining the notation used in the definition. (2)

[ii] You are investigating the survival times of two cancer patients who have just undergone chemotherapy with Cancer type A and B. You have recorded the following data for each patient:

$$Z_1 = \begin{cases} 0 \; if \; Non \; Smoker \\ 1 \; if \; Smoker \end{cases}$$

$$Z_2 = \begin{cases} 1 \; if \; having \; Cancer \; type \; B \\ 0 \; otherwise \end{cases}$$

You have decided to model the force of mortality at time t (measured in days since the treatment was provided) by using the above model (as in i.), and you have estimated the parameter values to be

$$\acute{\beta}_1 = 0.017$$

$$\acute{\beta}_2 = 0.026$$

Compare the force of mortality for a smoker who is having cancer type A with that of a smoker who is having cancer type B. (3)

[5]

Q. 2) Outline the key differences between deterministic models and stochastic models. Give an example each of deterministic and stochastic models. **[5]**

Q. 3) A specialist investment fund sells long term investment products for a term of ten years. However, in recent years, the fund managers have noticed that a large proportion of customers withdraw their investments prematurely before the maturity period. In order to devise a more appropriate investment strategy, the fund managers wish to understand the duration for which each policy holder is likely to remain invested. The fund managers have collected the following data:

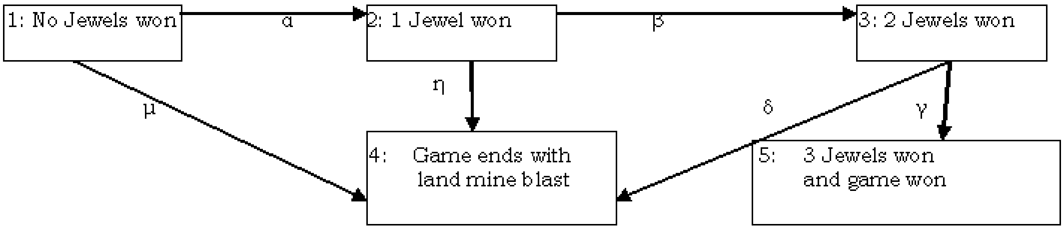
- List of unique customer ids' in-force at 31st March for each of the years from 2008-2012.
- Details of all exits from the fund between 2008 – 2012, including unique customer ids', date of exit and reason for exit (i.e. either maturity or withdrawal).
- Exact date of entry and unique customer ids' for all new policies sold since 2002.
- Due to a systems transformation, sales records are not available for policies sold prior to 2002. Nevertheless, the year of entry can be inferred from the unique customer id.

Define each of the following terms and explain whether or not the following types of censoring are present in the investigation described above:

i. Left censoring
ii. Random censoring
iii. Type II censoring
iv. Informative censoring
v. Interval censoring **[10]**

Q.4) In a computer game, the player needs to collect three jewels to win a stage. The game is programmed such that the level of difficulty increases with the number of jewels won. During the course of the game, there is a chance that the player may fail to counter a land mine; and if this happens, the player loses the game immediately.

The following continuous-time Markov chain model has been proposed for the progression of game:



The forces of transition are assumed to be independent of the previous state. The probability that a game in state i at exact time x is in state j at exact age x+t is denoted by ${}_tp_x^{ij}$.

[i] Derive from first principles an expression for $\frac{\delta}{\delta t}\,{}_tp_x^{24}$ (5)

[ii] Write down (without proof) expressions for $\frac{\delta}{\delta t}\,{}_tp_x^{22}$ and $\frac{\delta}{\delta t}\,{}_tp_x^{23}$ (3)

[iii] Hence or otherwise find ${}_tp_x^{22}$ explicitly. (3)
[11]

Q.5) Consider a game of tennis between two players Novak Djokovic and Andy Murray.

A game of tennis begins with the score being 0-0. If a player wins a point, his score would move from 0 to 15 and likewise from 15 to 30 and 30 to 40 upon winning subsequent points. If a player, who is already on 40, wins the next point, he wins the game except if the other player is on 40 at that time or is holding an “Advantage” at that time.

If both players are on 40 each, such situation is termed as “Deuce”. The player winning the first point from the state of Deuce is said to hold an “Advantage”. If the player holding the advantage wins the next point, he wins the game. However, if the next point is instead won by the other player, the game reverts to Deuce.

In a game, therefore, there are essentially 17 different states: 0-0, 15-0, 30-0, 40-0, 15-15, 30-15, 40-15, 0-15, 0-30, 0-40, 15-30, 15-40, Deuce, Advantage Djokovic, Advantage Murray, Game Djokovic, Game Murray.

You may note that from a modeling perspective:

- 30-30 is identical to Deuce
- 40-30 is identical to Advantage Djokovic
- 30-40 is identical to Advantage Murray

Let us assume that the probability of Djokovic winning a point is p, and that this probability is constant throughout the duration of the game. The probability of Murray winning a point may be denoted by q, where q = 1 – p.

Let $P_D(i,j)$ denote the probability that Djokovic wins the game given that his score is *i* and Murray's score is *j*.

[i] Calculate the probability $P_D(30,30)$ - the probability that Djokovic wins the game from a position where both players are on 30 each. (5)

[ii] Using your probability in (i) above or otherwise, calculate $P_D(40,30)$ and $P_D(30,40)$ (3)

[iii] Show that the probability for Djokovic winning the game, $P_D(0,0)$, is

$$P_D(0,0) = p^4 + 4p^4q + \frac{10p^4q^2}{1-2pq}$$

Hint: You may want to separately calculate the following probabilities and use them to determine the probability of Djokovic winning the game:

- the probability that Djokovic wins the game without losing a single point
- the probability that Djokovic wins the game losing only one point
- the probability that Djokovic wins the game after losing at least two points (7)

[15]

Q.6) A niche general insurance company specializes in certain corporate products. Sales fluctuate between two levels — "Low" and "High". It is experienced that introduction of new products has a positive impact on sales. However, the development of new products is a costly affair.

Introduction of new product in any year influences the sales in the following year. During the last quarter of the year, the company would develop and launch a new product so that the optimal impact of launching the new product is achieved in the following year. During the last quarter, the company can accurately forecast whether the sales for the current year would be "Low" or "High" since three-fourth of the year has already elapsed.

Introduction of a new product would cost the company by way development, promotion and marketing expenses. Such cost is estimated to be Rs 1 Crore for the year in which the product is developed and launched. Upon the launch of a new product, the probability of having ”High” sales the next year is 1/2 or 3/4, depending upon whether the current year’s sales are "Low" and "High". These probabilities go down to 1/4 or1/2 if a new product is not introduced during the current year. The company’s yearly profits (excluding the cost of developing new products if any) are Rs 10 Crores when sales are high but only Rs. 5 Crores when sales are low.

The CEO feels that the impact on incremental sale would be more if the new product is launched when sales are “Low”. Hence his proposal is to introduce new products when sales for the current year are “Low” while not launching new products when current year’s sales are “High”.

[i] Write down the (one-step) transition matrix for each of the following strategies: a) never introduce new products, b) introduce one new product every year c) introduce one new product as suggested by CEO depending on current year sales. (3)

[ii] Calculate the steady-state probabilities for each of the three cases in part (i). (6)

[iii] Find the long run expected average profit (allowing for costs of introducing a product) per year for each of the three strategies in part (i). Which of these strategies is best in terms of long run profitability? (4)

[13]

Q.7) [i] A life insurance company has obtained the graduated mortality rates for ages 41-60 shown in the table below. Carry out a test to determine whether the graduated rates may be considered smooth. Clearly bring out your conclusions and explain why these rates may be of a concern for a pricing actuary.

Age	Mortality rate	Age	Mortality rate
41	0.002495	51	0.008154
42	0.002894	52	0.007994
43	0.003360	53	0.007881
44	0.003892	54	0.007819
45	0.004488	55	0.007815
46	0.005146	56	0.007874
47	0.005864	57	0.008004
48	0.006641	58	0.008211
49	0.007473	59	0.008506
50	0.008356	60	0.008897

(6)

[ii] The life insurance company has carried out a series of additional tests to validate the graduated rates obtained. A summary of results from these tests is given below. For each test,

[a] State the purpose of the test; and

[b] Form appropriate conclusions and comment on the results of the test.

1 Chi-square test

Number of ages: 20

Degrees of freedom: 17

Value of test statistic: 14.87

2 Standardised deviations test

Interval	$(-\infty,-3)$	$(-3,-2)$	$(-2,-1)$	$(-1,0)$	$(0,1)$	$(1,2)$	$(2,3)$	$(3,\infty)$
Actual:	0	1	3	6	7	3	0	0
Expected	0.0	0.4	2.8	6.8	6.8	2.8	0.4	0.0

(8)
[14]

Q.8) [i] Write down the likelihood of observing d deaths, assuming that the number of deaths follows a Poisson model. Define all symbols you use. (1)

[ii] Use the likelihood function above to derive an expression for the parameters of the statistical distribution for the maximum likelihood estimate of the rate parameter, μ. (5)

[iii] The following data has been collected from a mortality investigation of 3 homogenous lines of business for the integer age x:

Business unit	**Exposed to Risk**	**Number of Deaths**
1	1814	182
2	1943	217
3	1732	180

Use the results from part (ii) above to determine:

[a] the best estimate value of the rate parameter, μ

[b] the 99.5^{th} percentile value of the rate parameter, μ

[c] Hence, calculate the best estimate and the 99.5^{th} percentile value of q_x for the above data under the Poisson model. (6)

[iv] Interpret these results in the context of mortality risk faced by a life insurance company that uses the best estimate of q_x calculated above to determine its mortality assumptions. (3)
[15]

Q.9) [i] State two advantages of using central exposed to risk in actuarial investigations as opposed to initial exposed to risk. (2)

An investigation is carried out in respect of mortality of married women over the period 1 October 2008 to 1 October 2012. The following data has been collected for four specific lives:

	Date of birth	**Wedding date**	**Notes**
Rita	1 October 1979	1 May 2006	Rita died on 1 January 2010.
Sita	1 September 1981	1 November 2008	
Nita	1 December 1979	1 February 2010	Nita got divorced on 1 November 2010.
Gita	1 April 1980	1 June 2011	

[ii] Calculate the contribution of each of the above four lives to the central and initial exposed to risk at age 30 last birthday. (8)

[iii] Hence, also determine the total central and initial exposed to risk. A typical approximation used in actuarial investigations of exposed to risk is:

$$E_X = E_X^c + \frac{1}{2}d_x$$

In light of the results of initial and central exposed to risk or otherwise, explain why this is not a good approximation for the data provided above. (2)

[12]

****