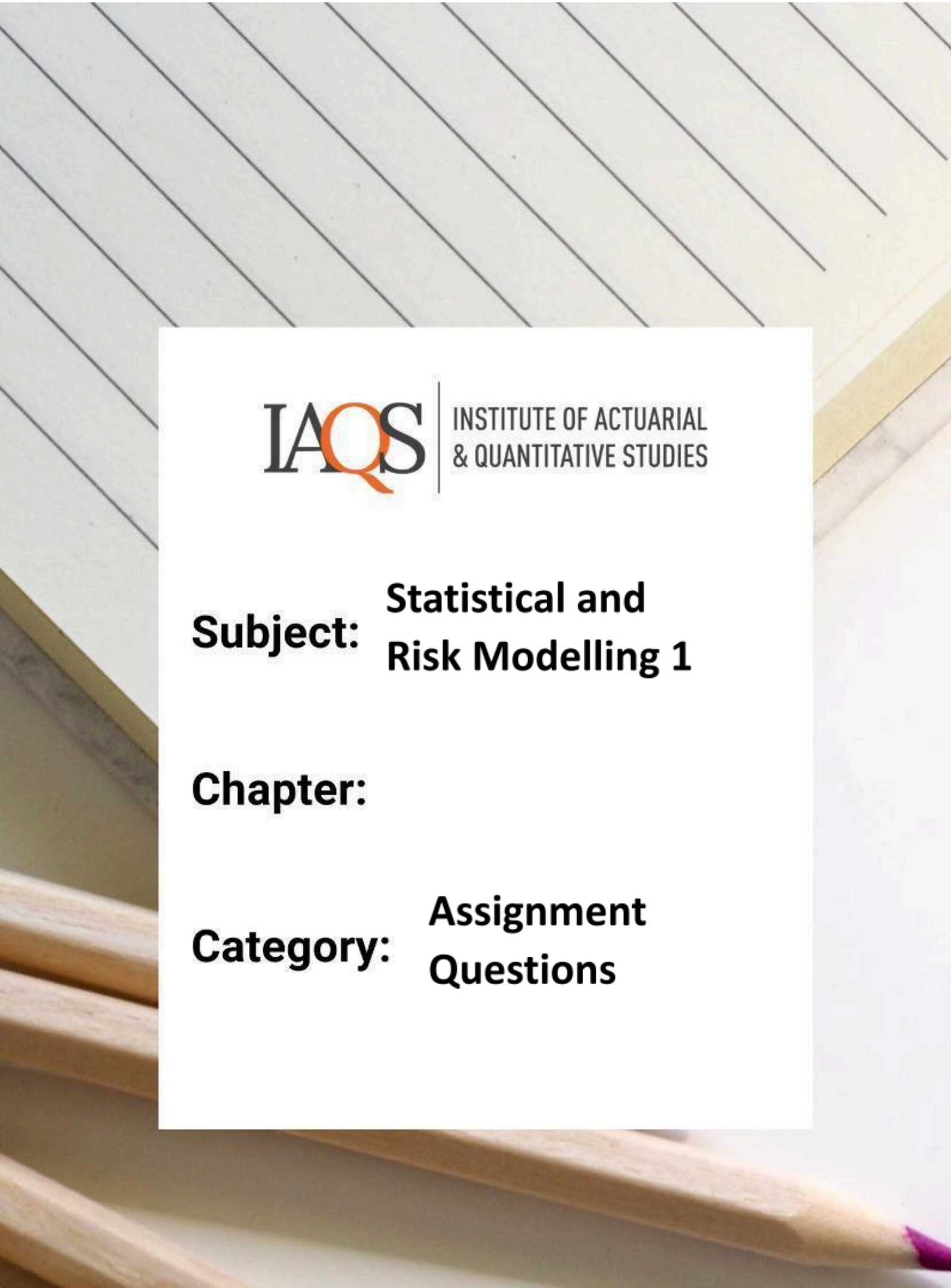

IAQS | INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Subject: Statistical and Risk Modelling 1

Chapter:

Category: Assignment Questions

1. It is assumed that the future lifetime of a life aged x is expressed as

$$T_x = K_x + S_x$$

where Kx = curtate life time, Sx is independent of Kx, has a uniform distribution in the interval (0,1).

Show that ${}_u q_x = u.q_x \qquad 0 < u < 1$

2. You are investigating mortality experience for the whole of the calendar years 1999 and 2000. A life being observed in this investigation attained exact age 54 on 1-06-2000 and died on 25-10-2000.

Calculate to the nearest number of weeks, the contribution of the above life to the:

a) Central exposed to risk and

b) Initial exposed to risk

for a population of lives aged 54 last birthday.

IAQS INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

3. With the success of a first gold medal in the Javelin throw event during the 2021 Olympics, the Bharat Olympics Committee (BOC) has started showing keen interest in the development of future javelin throwers for the country.

In search of the new talent, the Committee inducts new athletes for training if they throw a javelin for a distance of at least 50m in 3 attempts. Such inducted athletes participate in two events each year. Out of these inducted athletes, the committee considers an athlete as qualified for extensive Olympics level training if they threw a minimum distance of 80m in three attempts. The Central Government has decided to bear the cost of training support for such qualified athletes. The Government has asked BOC to assess the cost of providing such support and therefore, BOC wishes to know the average time it can expect to take for athletes to qualify for this extensive training.

The Committee has maintained a list of 23 athletes who have thrown at least 50m distance in 3 attempts by the start of 2014. The progress of those athletes has been recorded up to and including the last event of 2020. The following data records the number of events which had been held before the qualification of an athlete in this cohort:

Qualified 6, 8, 8, 9, 9, 9, 11, 11, 13, 13, 13
Injured and stopped participating 4, 5, 8, 11, 14

The remaining seven athletes were still participating in the events at the end of 2020.

i) Determine the median number of events taken to qualify for those athletes who qualified for Olympics level training during the period of observation.

ii) Calculate the Kaplan-Meier estimate of the survival function, S(t), for the "hazard" of qualifying, where t is the number of events since 1 January 2014.

iii) Using the result from (ii), estimate the median number of events to qualify for the athletes.

iv) Explain the reasons for the difference between the results in (i) and (iii) above.

4. i) State the precise mathematical formula representing the complete expectation of life, in terms of probability of survival. Explain in words what this represents.

You are given that the force of mortality, μ_x is constant 0.0325 at all ages. Calculate:

ii) the curtate expectation of life

iii) the probability that a life aged exactly 36 will survive to age 45.

iv) the exact age x representing the median of the life time T of a new born baby.

IAQS INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

5. i) State the age ranges over which Gompertz Law is an appropriate model for human mortality.

ii) Show that, under Gompertz Law, the probability of survival from age x to age x+t is equal to:

$$_tP_x = (g)^{c^x(c^t-1)}$$

Where g is defined as $logg =- B/logc$

6. A study has been conducted to investigate the effect of a newly invented drug on a group of patients who are suffering from cancer. The following proportional hazard regression model has been fitted to the mortality data of the group of patients.

$h_i(t) = h_0(t)xexp\{0.01(x_i - 30) + 0.2y_i - 0.05z_i\}$

where $h_i(t)$ denotes the hazard function for life i at duration t

$h_0(t)$; denotes the baseline hazard function at duration t

x_i;denotes the age at entry into the observation of life i

$y_i = 1$ if life i is a non − smoker, else 0 .

$z_i = 1$ if life i is a male , 0 if female

a) describe the class of lives to which the baseline hazard function applies.

b) using the model compare the survival function of a male smoker aged 30 at entry relative to a female smoker aged 40 at entry.

c) using the model compare the survival function of a male smoker aged 30 at entry relative to a male smoker aged 40 at entry.

7. An investigation studied the mortality of persons aged between exact ages 50 and 51 years. The investigation began on 1 January 2019 and ended on 31 December 2019. The following table gives details of 10 lives involved in the investigation.

Life	*Date of 50th birthday*	*Date of death*
1	1 February 2018	–
2	1 April 2018	1 October 2019
3	1 June 2018	–
4	1 September 2018	–
5	1 November 2018	15 March 2019
6	1 January 2019	–
7	1 May 2019	15 December 2019
8	1 July 2019	1 October 2019
9	1 August 2019	–
10	1 December 2019	–

Persons with no date of death given were still alive when the investigation ended.

i) Calculate the central exposed to risk using the data for the 10 lives in the sample.

ii) Calculate the Maximum Likelihood Estimate of the hazard of death at age 50 last birthday.

iii) Hence, or otherwise, estimate q_{50}.

8. An investigation was undertaken into mortality of patients suffering from a particular infection outbreak in a town who were treated in a hospital. A sample of patients was observed from the date of admission until they died or left the hospital or a period of 45 days elapsed. The following data relate to 13 patients treated in a hospital for infection.

Patient Number	Duration of Observation (days)	Reason for observation ceasing
1	5	Died
2	7	Died
3	14	Died
4	18	Discharged
5	25	Discharged
6	28	Died
7	30	Discharged
8	34	Discharged
9	35	Died
10	36	Discharged
11	45	Observation ended
12	45	Observation ended

IAQS INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

13	45	Observation ended

i) State the types of censoring that are present in the above investigation and explain where they occur?

ii) Comments on whether censoring likely to be informative in the investigation?

iii) Calculate the Kaplan-Meier estimate of the survivor function for the patients. State any assumption you have made.

At the end of the investigation, the hospital claims in a press conference that more than 80% of patient who responded to treatment for the initial two weeks are recovered from the infection and likely discharged from hospital.

iv) Comment on the hospitals claim in view of the above investigation?

IAQS INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

9. i) In a certain population, q_x = 0.3, calculate the value of mx (weighted average of force of mortality) assuming:

a) that deaths are uniformly distributed between the ages of x and x+1

b) a constant force of mortality between ages x and x+1

10. A life insurance company has used the Cox model for analyzing the persistency experience of its policy by the following parameters.

- Premium frequency (Annual, non-annual)
- Distributional channel (Online, Agency, Bancassurance)
- Method of premium payment (Direct debit, Cheque)

Using the following co-variates

F - Value 0 if frequency is annual and 1 if non-annual
D - Value 0,1 and 2 for Online, Agency and Bancassurance respectively
M - Value 0 for Direct debit and 1 for Cheque
The associated parameters are β_F, β_D and β_M respectively.

i) Why is the Cox model also called a proportional hazard model?

ii) Give the equation for the Cox model used in the analysis, defining any additional terms used.

iii) State the characteristics of the policy to which the baseline hazard applies.

The results showed that

- The hazard rate of lapse for Annual policy sold through Agency and paid by direct debit is 25% lower than hazard rate of non-annual policy sold through Agency and paid by cheque.

- Premiums paid through direct debit, annual policies of Agency channel had same hazard rate of lapse as non-annual policies of Online channel.

- The hazard rate of lapse for Annual policies of Online channel where premiums were paid through cheque was $3/4^{th}$ the hazard rate of lapse for annual policies of Bancassurance channel where premiums were paid through direct debt.

iv) Calculate the estimated values of the parameters β_F, β_D and β_M

IAQS INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

11. A study of the mortality of 12 laboratory bred insects was undertaken. The insects were observed from birth until either they died or the period of study ended, at which point these insects still alive were treated as censored.

The following table shows the Product-Limit estimate of the survival function, based on data from the 12 insects.

$t\ (weeks)$	$S(t)$
0 < t < 1	1.00
1 < t < 3	0.9167
3 < t < 6	0.7130
t > 6	0.4278

ii) Calculate the number of insects dying at durations 3 & 6 weeks.

iii) Calculate the number of insects whose history was censored.

12. An investigation was carried out for mortality of people aged between 50 and 60 years collecting data on the following covariates for each person:

X_1 Sex (a categorical variable with 0 = female, 1 = male)

X_2 Cigarette smoking (a categorical variable with 0 = non-smoker, 1 = smoker)

i) Sate the Gompertz Law.

In addition, data were collected on the age at death for persons who died during the period on investigation.

For analysis of the above data, it was decided to use a Gompertz hazard, $\lambda_x = Bc^x$ where age (x) is assumed as the duration since the start of the observation.

ii) Show that the substitution:

$$B = exp\left(\beta_0 + \beta_1 X_1 + \beta_2 X_2\right)$$

in the Gompertz model, leads to a proportional hazards model for this particular analysis.

Using the Gompertz hazard, the parameter estimates in the proportional hazards model were as follows:

Covariate	Sex	Smoker
Parameter estimate	β_1	β_2
Parameter	+0.30	+0.65

The parameter value of β_0 *is* -4.00 *and c is* $+1.05$

iii) Describe the characteristics of the person to whom the baseline hazard applies in this model.

iv) Calculate the estimated hazard for a female cigarette smoker aged 54 years.

v) Show that according to this model, a cigarette smoker at any age has a risk of death roughly equal to that of a non-smoker aged 13 years older.

13. An insurance company is carrying out mortality investigation of its term assurance portfolio. It records in-force policies using age label 'age y last birthday' as at 1st April. Information about the number of in-force policies is available for years 2015, 2016 and 2017. The number of deaths in financial year 2015 to 2017 as reported by claim

department grouped by age x nearest birthday on the date of death. No unreported claims are assumed.

The following data have been supplied for the investigation:

	Age			
	55	56	57	58
No. of deaths	1150	1380	1420	1780
No. of lives at 01.04.2017	18500	20000	15000	21200
No. of lives at 01.04.2016	20500	21100	20700	20500
No. of lives at 01.04.2015	20100	20000	19700	18500

i) Estimate force of mortality for lives with age label 56 and 57, state any assumptions made.

ii) Estimate initial mortality rates for lives in (i) using derived force of mortality, clearly indicating the age to which it applies to.

14. The Cox proportional hazard model it to be used to model the rate at which people are getting married. Assuming they get married within 3 years once they start looking for their life partner. In the fitted model, the hazard depends on time, t, since starting the search for a life partner. The covariates, their categories and the fitted parameters for each category are shown below:

Covariate	Possibility	Parameter
Profession	Service	0.3
	Business	0.5
	House Maker	0

Gender	Social Service	− 0.1
	Male	0
Location	Female	0.3
	Non Metro	0
Age Band	20 − 25	0.2
	25 − 30	0.7
	30 − 35	0.5
	35 − 40	0

IAQS INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

i) Defining clearly all the terms you use, write down an expression for the hazard function for this model.

ii) For a female social worker aged 37 living in Mumbai, who has been looking to get married since last 1 year, the probability for staying single for next 2 years is 0.3. Calculate the probability that a male aged 24 living in non-metropolitan city and doing business will stay single for next 2 years given that he is also looking for a partner since last 1 year.

15.

i) State two advantages of using central exposed to risk in actuarial investigations as opposed to initial exposed to risk.

An investigation is carried out in respect of mortality of married women over the period 1 October 2008to 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 four lives to central and initial exposed to risk at age 30 last birthday.

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

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

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