INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Subject: SRM1

Chapter: Unit-4

Category: Practice Questions

1. **CS2A A2022 Q1**

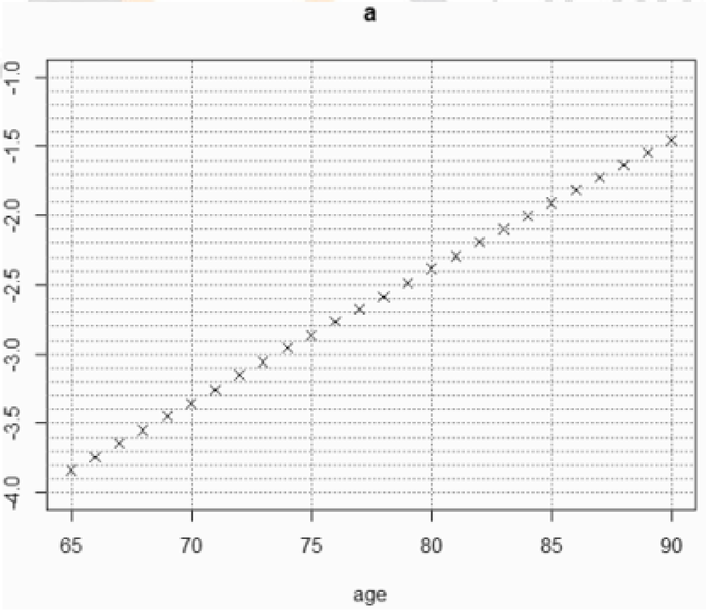
(i) Write down the two-factor Lee–Carter model, clearly defining each of the terms you use.

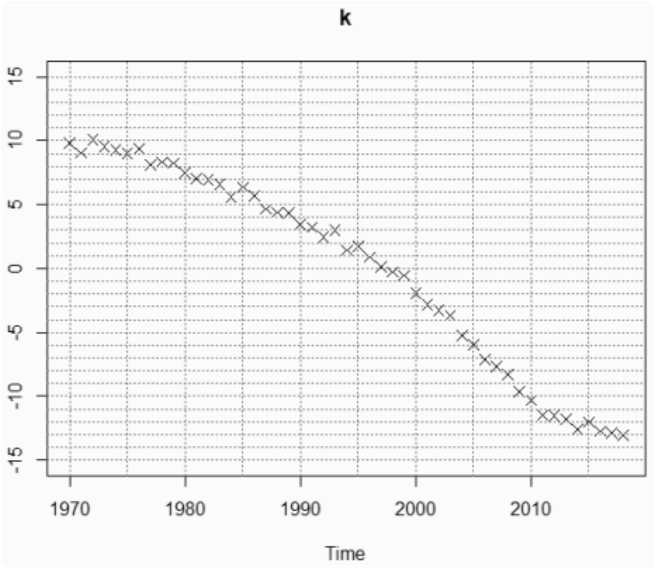
A statistician is using the two-factor Lee–Carter model to project future mortality rates and has fitted the model to a set of mortality data. The statistician has observed that the fitted forces of mortality do not vary regularly with each calendar year but vary more regularly with age. Therefore, the statistician suggests that, before projecting future mortality rates, the k parameters in the Lee–Carter model should be smoothed using penalised regression splines, but the a and b parameters should be kept as separate, unsmoothed parameters for each age.

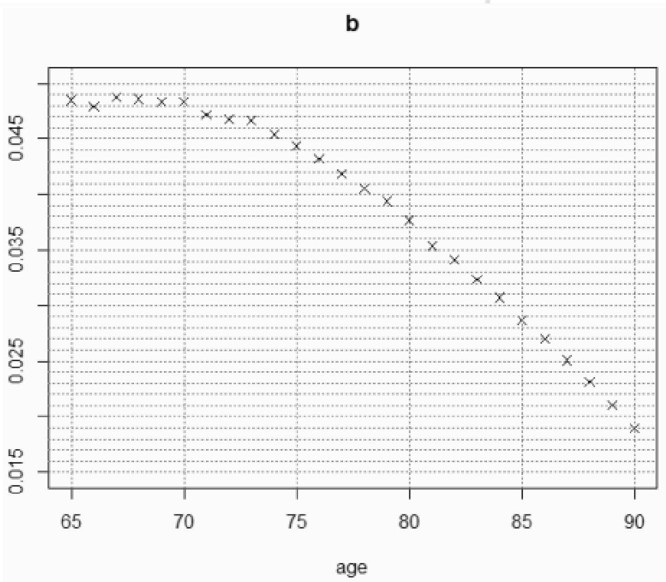
(ii) Discuss the statistician's suggestion by considering each of the three parameters of the Lee–Carter model, k, a and b, in turn.

2. **CS2A A2023 Q3**

The figure below shows three vectors ax, bx and kt, which were obtained by applying the two-factor Lee–Carter model to mortality data derived from the period 1970 to 2018 for the age range 65 to 90 years.







(i) Discuss what each of these three vectors suggest about the characteristics of the mortality rates in the data.

(ii) Suggest a socio-economic reason for your conclusion about vector b in part (i) above.

(iii) Calculate $m_{70,\ 2018}$ using the output from the model (reading the values from the plots in the figure).

On analysing the past data over a larger age range (from 20 to 90 years) it was determined that there was an increase in rates for 20–30 year olds in the 1970s due to an illness that almost exclusively affected the younger population. The higher mortality rates only lasted until 1980 when a cure was found and introduced to the population.

(v) Explain how this cohort effect may manifest itself in the output from the two-factor Lee–Carter model, fitted to ages 20–90 and calendar years 1970–2018, in terms of a, b and k.

3. IAQS PAST PAPER

For the senior citizen population in a country, a statistician has decided to use the two-factor Lee–Carter model to project future mortality rates and has fitted the model to a set of mortality data.

i) Write down the two-factor Lee–Carter model, clearly defining each of the terms you use. Also list out the constraints that are normally imposed, in order for the model to be uniquely specified.

ii) The parameters ax and bx of the model for ages between 60 to 80 (inclusive) are expressed using linear functions as below.

ax = 0.105x - 10.95

bx = -0.004x + 0.48

Further, it is assumed that the factor kt decreases linearly from 2.75 at time 0 to -1.25 at time 40.

Estimate the central mortality rate for ages 60 and 70 and at times 0, 10, 20, 30 and 40. (6)

iii) From the results in ii above, comment on the trend in central mortality rate with time for ages 60 and 70.

iv) Describe the advantages and disadvantages of Lee-Carter Model.

Solution:

i)

The two-factor Lee-Carter model may be written:

$\ln m_{x,t} = a_x + b_x * k_t + \varepsilon_{x,t}$

where:

$m_{x,t}$ is the central mortality rate at age x in year t

[0.5]

a_x describes the general shape of mortality at age x or a_x is the mean of the time- averaged logarithms of the central mortality rate at age x.

[0.5]

b_x measures the change in the rates in response to an underlying time trend in the level of mortality of k_t.

[0.5]

k_t reflects the effect of the time trend on mortality at time t, and

[0.5]

$\varepsilon_{x,t}$ are independently distributed normal random variables with means of zero and some variance to be estimated.

[0.5]

The usual constraints imposed in the Lee-Carter Model are that $\Sigma_x b_x=1$

and

$\Sigma_t k_t=0$

[0.5]

[3]

ii)

Ignoring error term $\varepsilon_{x,t}$, the mortality rate at age x in projection year t is:
$m_{x,t} = \exp(a_x + b_x * k_t)$
[0.5]

The time trend factor k_t is assumed to decrease linearly from 2.75 at time 0 to -1.25 at time 40. Hence, can be expressed as
$k_t = 2.75 – \_0.1*t$
[0.5]

Further, parameters a_x and b_x of the model for ages between 60 to 80 (inclusive) are expressed using linear functions as below.

$a_x = 0.105x - 10.95$

&

$b_x = -0.004x + 0.48$

X	ax	bx
60	-4.65	0.24
70	-3.6	0.2

[1]

From the above,
$m_{60,t} = \exp(a_{60} + b_{60}*k_t)$
$= \exp(-4.65 + 0.24 *(2.75 – 0.1t))$
$= \exp(-4.65 + 0.66 – 0.024t)$
$= \exp(-3.99 – 0.024t)$
[1]

$m_{60,0} = \exp(-3.99) = 0.0185$
$m_{60,10} = \exp(-3.99 – 0.24) = \exp(-4.23) = 0.014552$
$m_{60,20} = \exp(-3.99 – 0.48) = \exp(-4.47) = 0.011447$

m60,30 = exp (-3.99 – 0.72) = exp (-4.71) = 0.009005
m60,40 = exp (-3.99 – 0.96) = exp (-4.95) = 0.007083
[1]

Similarly,
m70,t = exp (a70 + b70*kt)
= exp(-3.6 + 0.2 *(2.75 – 0.1t))
= exp (-3.6 + 0.55 – 0.02t)
= exp (-3.05 – 0.02t)
[1]

m70,0 = exp (-3.05) = 0.047359
m70,10 = exp (-3.05 – 0.2) = exp (-3.25) = 0.038774
m70,20 = exp (-3.05 – 0.4) = exp (-3.45) = 0.031746
m70,30 = exp (-3.05 – 0.6) = exp (-3.65) = 0.025991
m70,40 = exp (-3.05 – 0.8) = exp (-3.85) = 0.02128
[1]

[6]

Alternatively, Ignoring error term εx,t, the mortality rate at age x in projection year t is:
mx,t = exp (ax + bx * kt)
[0.5]

The time trend factor kt is assumed to decrease linearly from 2.75 at time 0 to -1.25 at time 40.
Hence, can be expressed as
kt = 2.75 – 0.1*t
[0.5]

Further, parameters ax and bx of the model for ages between 60 to 80 (inclusive) are expressed using linear functions as below.

ax = 0.105x - 10.95

&

bx = -0.004x + 0.48

mx,t = exp (0.105x - 10.95 + (-0.004x + 0.48)* (2.75 – 0.1*t))
= exp (0.105x - 10.95 -0.004*2.75x + 0.004*0.1*x*t + 0.48*2.75 -0.048*t)
= exp (0.105x - 10.95 – 0.011x + 1.32 – 0.048*t +0.0004*x*t)
= exp (0.094x – 9.63 - 0.048*t +0.0004*x*t)
[2]

m60,0 = exp (-3.99) = 0.0185
m60,10 = exp (-4.23) = 0.014552
m60,20 = exp (-4.47) = 0.011447

$m_{60,30} = \exp(-4.71) = 0.009005$
$m_{60,40} = \exp(-4.95) = 0.007083$
[1.5]

$m_{70,0} = \exp(-3.05) = 0.047359$
$m_{70,10} = \exp(-3.25) = 0.038774$
$m_{70,20} = \exp(-3.45) = 0.031746$
$m_{70,30} = \exp(-3.65) = 0.025991$
$m_{70,40} = \exp(-3.85) = 0.02128$
[1.5]

[6]

iii) For every year that we project into the future, the mortality rate $m_{60,t}$ is multiplied by a factor of $e^{-0.024}$ (or 0.9763), *i.e.* it decreases by approximately 2.4% *p.a.*

[1]

For $m_{70,t}$, we multiply by a factor of $e^{-0.02}$ (or 0.9802) for each year that we project into the future and hence the values of $m_{70,t}$ decrease by approximately 2% *p.a.*

The percentage reduction in $m_{70,t}$ is smaller than that for $m_{60,t}$ since $b_{70} < b_{60}$.

[1]

[2]

iv) Future estimates of mortality at different ages are heavily dependent on the original estimates of the parameters ax and bx, which are assumed to remain constant into the future. These parameters are estimated from past data, and will incorporate any roughness contained in the data. In particular, they may be distorted by past period events which might affect different ages to different degrees.

• If the estimated bx values show variability from age to age, it is possible for the forecast age- specific mortality rates to 'cross over' (such that, for example, projected rates may increase with age at one duration, but decrease with age at the next).

• There is a tendency for Lee-Carter forecasts to become increasingly rough over time.

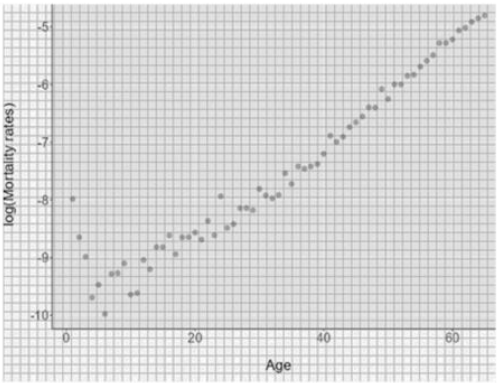
• The model assumes that the underlying rates of mortality change are constant over time across all ages, when there is empirical evidence that this is not so.

• The Lee-Carter model does not include a cohort term, whereas there is evidence from some countries that certain cohorts exhibit higher mortality improvements than others.

• Unless observed rates are used for the forecasting, it can produce 'jump-off' effects (ie an implausible jump between the most recent observed mortality rate and the forecast for the first future period).

4. CS2A September 2024 Q7

The graph below shows the crude mortality rates from age 1 to 65 on a logarithmic scale in a population



You have recently joined an insurance company as a risk analyst and your team is interested in using this data to project mortality rates into ages above 65.

One of your colleagues proposes a model of the form:

$$y_x = \exp(\alpha + \beta x + \varepsilon_x) \text{ with } \varepsilon_x \sim N(0, \sigma^2)$$

where x represents age, yx represents the mortality rate at age x, and α and β are two unknown parameters that can be estimated.

(i) Comment on the suitability of Model 1 for projecting mortality rates into ages above 65 years.

After fitting Model 1, your line manager suggests exploring alternative models in the following form:

$$y_x = \exp\left(\sum_k f_k(x)a_k + \varepsilon_x\right)$$

where fk () are known functions of your choice and the ak are unknown coefficients that can be estimated.

(ii) Discuss two variants of Model 2 that can be used to tackle the shortcomings of Model 1 stating the advantages and limitations of each.

Your company has recently acquired new mortality data, separate from the graph above. These data include aggregated death counts and exposed-to-risk by age x and calendar year t.

Your company aims to use these data for forecasting future mortality rates. Your manager recently participated in a mortality modelling seminar where the presenter successfully used the following simplified Lee–Carter model to forecast mortality rates:

$$y_{x,t} = \exp(a_x + k_t + \varepsilon_{x,t}) \text{ with } \varepsilon_{x,t} \sim N(0, \sigma^2)$$

where a_x and k_t are unknown parameters that can be estimated.

Based on that experience, your line manager suggests that the team should consider Model 3.

(iii) Demonstrate that Model 3 is not identifiable.

(iv) Describe how you would achieve identifiability when fitting Model 3.

(v) Comment on the suitability of Model 3.