# Institute of Actuaries of India

# Subject CT4 – Models

# May 2015 Examinations

# INDICATIVE SOLUTIONS

**Solution 1 :**

**(i)**

$S_x$ is the survival function of $T_x$

For e.g. $S_x(t)$ represents the probability that a person currently aged x survives for a period 't' years from now.

Let the probability density function (pdf) be defined by $f_{(x)}(t)$.

Then $f_{(x)}(t) = \frac{d}{dt} F_x(t)$

$f_{(x)}(t) = \frac{d}{dt} P[T_x \leq t]$

$= \lim_{h \to 0+} \frac{1}{h} * \frac{(P[T_x \leq t+h] - (P[T_x \leq t])}{}$

$= \lim_{h \to 0+} \frac{P[T \ \ \leq x+t+h | T > x] - (P[T \ \ \leq x+t | T > x]}{h}$

$= \lim_{h \to 0+} \frac{P[T \ \ \leq x+t+h] - P[T \leq x] - (P[T \ \ \leq x+t] - P[T \leq x])}{S_x * h}$

$= \lim_{h \to 0+} \frac{P[T \ \ \leq x+t+h] - (P[T \ \ \leq x+t])}{S_x * h}$

Multiplying and dividing by S(x+t)

$f_{(x)}(t) = \frac{S(x+t)}{S(x)} * \lim_{h \to 0+} \frac{1}{h} * \frac{P[T \ \ \leq x+t+h] - (P[T \ \ \leq x+t])}{S(x+t)}$

$= S_x(t) * \lim_{h \to 0+} \frac{1}{h} * P[T_x \leq t+h | \ T > x+t]$

$= S_x(t) * \mu_{x+t}$

The density function of Tx represents the probability of surviving till the time t and dying in the moment between t and t+h. **[3]**

**(ii)** The expectation of life at any particular age say 'x' refers to the expected future lifetime after age. The symbol $e_x$ refers to the curtate expectation of life and is defined as

$e_x = E[K_{x.}]$ where $K_x$ is the curtate future lifetime of a life aged x.

$e_x = \sum_{k=0}^{w-x} k * Sx(k) * q_{x+k}$ **[2]**

**(iii)** $e_x = {}_k p_x * (k + e_{x+k})$ where k is the period upto which a person currently aged 'x' has survived when the expected future lifetime is $e_{x+k}$

Hence $e_0 = {}_{40} p_0 * (40 + e_{40}) = {}_{40} p_0 * (40 + e_{40})$

$e_{40} = e_0 / {}_{40}p_0 - 40 = 60 / {}_{40}p_0 - 40$ which will be equal to 20 only if ${}_{40}p_0 = 1$, which is not possible as there will be some deaths between age 0 to 40.

Similarly

$e_{55} = e_0 / {}_{55}p_0 - 55 = 60 / {}_{55}p_0 - 55$ which will be equal to 5 only if ${}_{55}p_0 = 1$, which will again not be possible as there will be some deaths between age 0 to 55. **[2]**

**[7 Marks]**

**Solution 2 :**

**(i)**

Benefits of modeling in actuarial work

- Systems with long time frames such as the operation of a pension fund can be studied in compressed time.
- Different future policies or possible actions can be compared to see which best suits the requirements or constraints of a user.
- Complex situations can be studied.
- Modeling may be the only practicable approach for certain actuarial problems.

**[2]**

**(ii)**

A model is described as stochastic if it allows for the random variation in at least one input variable. Often the output from a stochastic model is in the form of many simulated possible outcomes of a process, so distributions can be studied.

A deterministic model can be thought of as a special case of a stochastic model where only a single outcome from the underlying random processes is considered.

Sometimes stochastic models have analytical/closed form solutions, such that simulation is not required, but they are still stochastic as they allow for factors to be random variables.

if the distribution of possible outcomes is required then stochastic modeling would be needed, or if only interested in a single scenario then deterministic.

Budget and time available stochastic modeling can be considerably more expensive and time consuming.

**[2]**

**(iii)**

The following factors may favor a stochastic approach:

- The regulator may require a stochastic approach.
- Extent of non-linear variation for example existence of options or guarantees.
- Skewness of distribution of underlying variables, such as cost of storm claims.
- Interaction between variables, such as lapse rates with investment performance.

The following may favor a deterministic approach:

- Lack of credible historic data on which to fit distribution of a variable.
- If accuracy of result is not paramount, for example if a simple model with deliberately cautious assumptions is chosen so as not to under estimate costs.

**[3]**

**(iv)**

A deterministic result on best estimate assumptions could be compared with the mean and median outcomes from a stochastic approach.

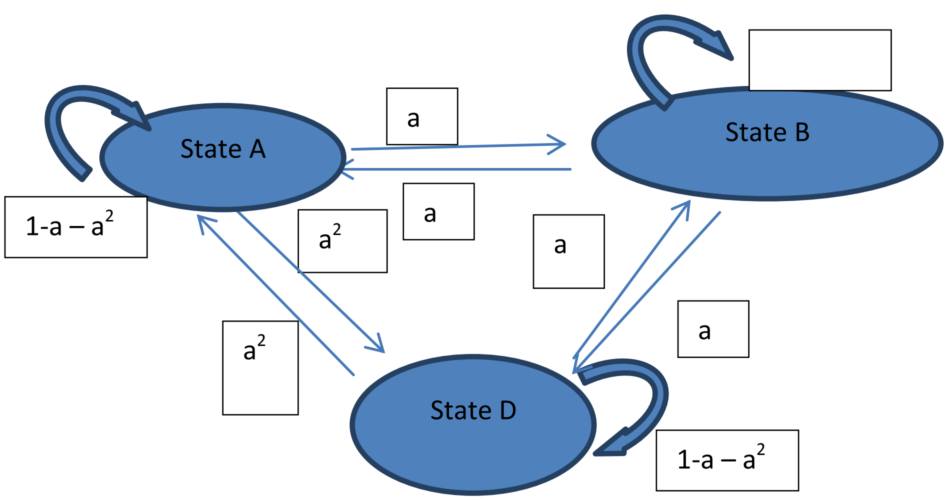
A deterministic model may also be used to calculate the expected or median outcome, with a stochastic approach being used to estimate the volatility around the central outcome. **[2]**

**[9 Marks]**

**Solution 3 :**

**(i)**

Transition graph given below




**[2]**

**(ii)** Transition probabilities must lie in [0,1]. Thus we need $a >= 0$, $1 - 2a >= 0$ and $1 - a - a^2 >= 0$

The solution of the quadratic is the interval $[-1/2 - \sqrt{5}/2, -1/2 + \sqrt{5}/2]$ , so all conditions are satisfied simultaneously fora€ [0,1/2]

**[2]**

**(iii)** The chain is both irreducible, as every state can be reached from every other State, and aperiodic, as the chain may remain at its current state for all A,B, D

**[1]**

**(iv)** From the result in (iii), a stationary probability distribution exists and it is unique. Let $\pi = (\pi_A, \pi_B, \pi_D)$ denote the stationary distribution. Then, can be determined by solving $\pi P = \pi$

For $a = 0.2$, the transition matrix becomes

$$\mathbf{P} = \begin{vmatrix} 0.76 & 0.2 & 0.04 \\ 0.2 & 0.6 & 0.2 \\ 0.04 & 0.2 & 0.76 \end{vmatrix}$$

So that the system $\pi P = \pi$ reads

$0.76\,\pi_A + 0.2\,\pi_B + 0.04\,\pi_D = \pi_A$ (1)

$0.2\,\pi_A + 0.6\,\pi_B + 0.2\,\pi_D = \pi_B$

$0.04\,\pi_A + 0.2\,\pi_B + 0.76\,\pi_D = \pi_D$ (2)

Discard the second of these equations and use also that the stationary probabilities must also satisfy

$\pi_A + \pi_B + \pi_D = 1$ (3)

Subtracting (2) from (1) gives $\pi_A = \pi_D$.

Substituting into (1) we obtain $\pi_A = \pi_B$,

thus (3) gives that $\pi_A = \pi_B = \pi_D = 1/3$

The proportion of debt who are in state D in the long run is 1/3. **[3]**

**(v)** The second order transition matrix is

$$\mathbf{P} * \mathbf{P} = \begin{vmatrix} 0.76 & 0.2 & 0.04 \\ 0.2 & 0.6 & 0.2 \\ 0.04 & 0.2 & 0.76 \end{vmatrix} * \begin{vmatrix} 0.76 & 0.2 & 0.04 \\ 0.2 & 0.6 & 0.2 \\ 0.04 & 0.2 & 0.76 \end{vmatrix} = \begin{vmatrix} .6192 & 0.28 & 0.1008 \\ 0.28 & 0.44 & 0.28 \\ 0.1008 & 0.28 & 0.6192 \end{vmatrix}$$

The relevant entries are those in the last column, so that the answers are:
(a) 0.1008
(b) 0.28
(c) 0.6192. **[3]**

**[11 Marks]**

**Solution 4 :**

**(i)**

Assume that each life dies with probability $q_x$ and survives with probability 1- $q_{x.}$

Then the number of deaths 'D' has a binomial distribution with parameters N and $q_{x.}$

Since the lives are independent, the probability that'd' individuals will die during the year and

'N–d' individuals will not die, is

$q_x^d * (1\text{-}q_x)^{N\text{-}d}$. The 'd' deaths can occur in $\frac{N!}{d!*(N-d)!}$ ways.

Hence P (D=d) = $\frac{N!}{d!*(N-d)!} * q_x^d * (1\text{-}q_x)^{N\text{-}d}$ **[2]**

**(ii)**

As the deaths follow a Binomial model the expected number of death is $Nq_{70}$, where 'N' is the number of annuitants being observed.

The observed number of deaths is an estimator of $q_{70}$ i.e $\hat{q}_{70}$ = d/N = 1,820/ 10,000 = 0.1820

As N is large, we can apply the Normal approximation to the Binomial model and hence the distribution of estimator $\hat{q}_{70}$ has a Normal distribution with mean $q_{70}$ and variance $q_{70}$* (1-$q_{70}$)/N .

The 95% confidence interval of $q_{70}$ is ($\hat{q}_{70} \pm 1.96 * \sqrt{\hat{q}70 * (1 - \hat{q}70)/\text{N}}$)

= (0.1820 ± 1.96* ($\sqrt{0.1820}( 1 - 0.1820)$ / 10,000))

= (0.17444, 0.18956) **[3]**

**(iii)** The model requires the employees to be observed from exact age x to x+1 for a year for the rate of mortality to be estimated. The following issues would arise when using this method in a company

- Employees would enter at different points in the age interval [ x, x+1] and hence would not be observed for the period of 1 full year
- Employees would exit at different points in the age interval [ x, x+1] and hence would not be observed for the period of 1 full year
- There may be decrements other than death during the observation period

Similarly there would be increments as well **[2]**

**(iv)** An assumption about the distribution of death needs to be made. One assumption could be that deaths occur uniformly in the age interval [x,x+1].

Alternatively, the Poisson model which does not require lives to be observed for the entire duration of 1 year can be used. **[2]**

**[9 Marks]**

**Solution 5 :**

**(i)**

$P'_{0,0}(t) = \mu P_{0,1}(t) - \lambda P_{0,0}(t)$, or a more general form such as $P'_{0,0}(t) = \sum P_{0,k}(t)\, \sigma_{k,0}$ **[1]**

**(ii)**

Since $P0,1(t) = 1 - P0,0(t)$,
we have $P0',0(t) = \mu(1 - P0,0(t)) - \lambda\, P_{0,0}(t)$. Any solution method will do,

e.g. $d/dt[e^{(\lambda+\mu)t} P_{0,0}(t)] = \mu e^{(\lambda+\mu)t}$, solved by $P0,0(t) = \frac{\mu}{\lambda+\mu} + C\, e^{-(\lambda+\mu)t}$ with $C$ being determined by the fact that $P_{0,0}(0) = 1$. **[3]**

**(iii)** $\mathbf{E}_0 O_t = \mathbf{E}_0 \int_0^t I s ds = \int_0^t \mathbf{E} 0 I s ds = \int_0^t P0,0(s) \mathrm{ds}$

$= \frac{\mu}{\lambda+\mu} t + \frac{\lambda}{(\lambda+\mu)(\lambda+\mu)} * (1 - \exp -(\lambda + \mu)\, \mathrm{t}\,)$ **[2]**

**(iv)** Since the process must be in state 0 or state 1 at all times, the solution is

just $t - \mathbf{E}_0 \mathrm{O}_t = \frac{\lambda}{\lambda+\mu} t - \frac{\lambda}{(\lambda+\mu)(\lambda+\mu)} * (1 - \exp -(\lambda + \mu)\, \mathrm{t}\,)$ **[1]**

**(v)** **a)** Assuming a member who is initially healthy, expected outgoings (Including expenses) by time *t* and expected income by time *t*, arerespectively

$$ct+b*(\frac{\lambda}{\lambda+\mu}t - \frac{\lambda}{(\lambda+\mu)(\lambda+\mu)} * (1- \exp -(\lambda+\mu)t))$$

and

$$a(\frac{\mu}{\lambda+\mu}t + \frac{\lambda}{(\lambda+\mu)(\lambda+\mu)} * (1- \exp -(\lambda+\mu)t))$$

In the long run, then, as $t\rightarrow\infty$, we require a $\mu$= b $\lambda$+ c($\lambda$+ $\mu$) to Break even.

**b)** The assumptions required are that the rate of becoming ill and rate of recovery from illness are constant.

**c)** This will certainly not be true of any individual member but, if membership is large and the age and health profiles of the members are constant by virtue of a constant influx of new members, it may be a reasonable approximation. **[3]**

**[10 Marks]**

**Solution 6 :**

**(i)**

Rate interval is defined as 'The period of one year during which a life's recorded age remains the same'.

As per the given data the age changes on a policy anniversary and hence it is a policy year rate interval. **[2]**

**(ii)** Let $P_{x,t}$ be the "number of policies in force aged x last birthday at the previous policy anniversary" at time t.

Then a consistent exposed to risk would be

$E_x^c = \int_0^1 P_{x.t}$dt .

Assuming that policy anniversaries are uniformly distributed across the financial year $E_x^c$ can be approximated as

$E_x^c$ = ½ [$P_{x,0}$ + $P_{x,1}$]

*However* $P_{x.t}$ is not available directly from the data. The values of $P_{x.t}*$, the number of policies in force aged 'x' last birthday is available. A person aged x last birthday at the previous policy anniversary at time't' would be aged between x and x+1 last birthday at time 't'.

Assuming that birthdays are uniformly distributed across the policy year, half of these lives would be aged 'x' and half would be aged 'x+1'. Hence

$P_{x,t} = ½[\ P_{x,t}{}^{*} + P_{x+1,t}{}^{*}]$

Substituting this in the equation for $E_x^c$

$E_x^c$= ½ [ ½ [$P_{x,0}{}^{*} + P_{x+1,0}{}^{*}$]+½ [$P_{x,1}{}^{*} + P_{x+1,1}{}^{*}$]] **[3]**

**(iii)** $\mu_{45} = h_{45}/E_{45}^c$

$E_{45}^c$= ½ $P_{45,0} + \sum_{t=1}^{t=1} P_{45,t}$ + ½ $P_{45,2}$

= ½[½ [$P_{45,0}{}^{*} + P_{45+1,0}{}^{*}$]] + ½ [$P_{45,1}{}^{*} + P_{45+1,1}{}^{*}$] + ½ [½ [$P_{45,2}{}^{*} + P_{45+1,2}{}^{*}$]

= ½ [½[29.947+ 29.120]] + ½ [30.479+30,103] + ½ [½[30,253+ 30.437] ]

= 60,230

$h_{45}$ = 1321

$\mu_{45}$ = 1321/ 60230 = 0.0219

$\mu_{45}$ represents mortality at age 44.5 as it is a policy year rate interval and assuming birthdays are uniformly distributed over the policy year and the force of mortality is constant over age 45. **[4]**

**(iv)** The assumption that policy anniversaries are uniformly distributed over the financial year is no longer relevant and hence the results would not be appropriate.

Further the assumption that birthdays are uniformly distributed across the policy year may not be appropriate if customers choose to buy policies close to their birthdays to avoid increase in premium rates. **[2]**

**(v)** The estimation of $q_x$ requires the initial exposed to risk. As the decrement being assessed is morbidity and policies continue to be covered for the risk and hence continue to contribute to exposed-to-risk even after a claim, there is no difference between Initial exposed to risk and central exposed to risk. **[2]**

**[13 Marks]**

**Solution 7 :**

**(i)**

The hypothesis requires the expected claims to be calculated based on the observed mortality experience in the last year

E(A_LY) = observed mort rate (LY) * Exp_risk_CY = Deaths_LY/Exp_risk_LY * Exp_risk_CY

This requires the exposed to risk at each age. As the exposed to risk is not directly available the expected deaths based on actual experience in last year can be derived as
E(A_LY) = Deaths_LY/(Exp_risk_LY * std_mort_rate) *(Exp_risk_CY * std_mor_rate)

= Deaths_LY/Exp_deaths_LY * Exp_deaths_CY

| Age | Deaths_LY | Exp_deaths_LY | Exp_deaths_CY | E (A_LY) E | Act_claims A | $X^2$ = (A-E)^2/E |
|---|---|---|---|---|---|---|
| **25** | 808 | 810 | 724 | 722 | 719 | 0.01 |
| **30** | 1,851 | 1,708 | 1,433 | 1,553 | 1,322 | 34.46 |
| **35** | 1,400 | 1,084 | 444 | 573 | 500 | 9.28 |
| **40** | 1,562 | 1,705 | 1,397 | 1,280 | 1,207 | 4.15 |
| **45** | 1,366 | 1,572 | 1,465 | 1,273 | 1,177 | 7.24 |
| **50** | 1,296 | 1,643 | 1,209 | 954 | 905 | 2.48 |
| **55** | 2,200 | 2,911 | 2,436 | 1,841 | 1,798 | 1.00 |

There are 7 ages. Hence the degrees of freedom is 7.

The upper 95% for $X^2_7$ for 7 degrees of freedom is 15.5. The observed value is 58.6 which is much higher than this, hence the hypothesis is rejected. **[8]**

**(ii)** Even if the mortality underlying the portfolio has not changed from last year, direct comparison with the last year's experience is not appropriate. As the observed deaths in last year are estimators of the underlying mortality, the underlying mortality represented either as a proportion of the standard mortality table or graduated mortality rates of the observed experience should be used for comparison of this year's experience. **[2]**

**(iii)**

The crude estimates of mortality based on the observed claims will progress more or less roughly, mainly due to random sampling errors and the individual rates having been estimated independently. In order to smooth the rates graduation is needed. **[2]**

**[12 Marks]**

**Solution 8 :**

**(i)**

Informative censoring means that the censored lives have a different experience from the lives that are not censored. Non-informative censoring means that the censoring has no effect on the experience of the censored lives.

Examples of informative censoring are

- Withdrawal from life insurance policies of better than average mortality lives
- Ill-health retirements from pension schemes because these are likely to be in worse than average health than the continuing members

Examples of non-informative censoring are

- End of investigation period
- Unit-linked policyholders surrendering the policy in the year of good investment returns **[2]**

**(ii)**

Old Medicine

$$\hat{S}_{KM}(t) = \begin{cases} 1 \quad \text{for } 0 \le t < 3 \\ 9/10 \text{ for } 3 \le t < 5 \\ 9/10 * 7/9 \text{ for } 5 \le t < 6 \\ 9/10 * 7/9 * 6/7 \text{ for } 6 \le t < 8 \\ 9/10 * 7/9 * 6/7 * 4/5 \text{ for } 8 \le t < 12 \end{cases}$$

$\hat{S}_{KM}(t) =$

$$\begin{cases} 1 & \text{for } 0 \le t < 3 \\ 0.9 & \text{for } 3 \le t < 5 \\ 0.7 & \text{for } 5 \le t < 6 \\ 0.6 & \text{for } 6 \le t < 8 \\ 0.48 & \text{for } 8 \le t < 12 \end{cases}$$

New Medicine

$\hat{S}_{KM}(t)$ =

$$\begin{cases} 1 & \text{for } 0 \le t < 8 \\ 7/8 & \text{for } 8 \le t < 10 \\ 7/8 * 5/6 & \text{for } 10 \le t < 12 \end{cases}$$

=

$$\begin{cases} 1 & \text{for } 0 \le t < 8 \\ 0.875 & \text{for } 8 \le t < 10 \\ 0.73 & \text{for } 10 \le t < 12 \end{cases}$$

The new medicine is more effective as the survival probabilities are higher. **[6]**

**(iii)** The probability of survival is given by $\hat{S}_{KM}(9)$ (New medicine) = 0.875 **[1]**

**[9 Marks]**

**Solution 9 :**

**(i)**

**a)** Poisson process

A Poisson process with rate λ is an integer-valued process $N_t$, $t \geq 0$ with the following properties:
$N_0 = 0$;
$N_t$ has independent increments;
$N_t$ has stationary increments, each having a Poisson distribution, i.e.

$P[N_t - N_s = n] = [\lambda(t-s)]^n e^{-\lambda(t-s)}] / n!$ , s<t,n = 0,1,2,…

**b)** Compound Poisson Process

Let $N_t$be a Poisson process, $t \geq 0$ and let $Y1$, $Y2$, …,$Yj$, …, be asequence of i.i.d. random variables. Then a compound Poisson processis defined by

$X_t = \sum_{j=1}^{Nt} Yj \quad t>=0$

**c)** Thining of poisson process

When the events in a Poisson process are of different types, each type occurring at random with a certain probability, the events of a particular type form a *thinned* process. The thinned process is also a Poisson process, with rate equal to the original rate multiplied by the probability for the type of event

**d)** Markov Jump process
A continuous-time Markov process $X$ ,$t$0 $t$ with a discrete state space S is called a Markov jump process. **[4]**

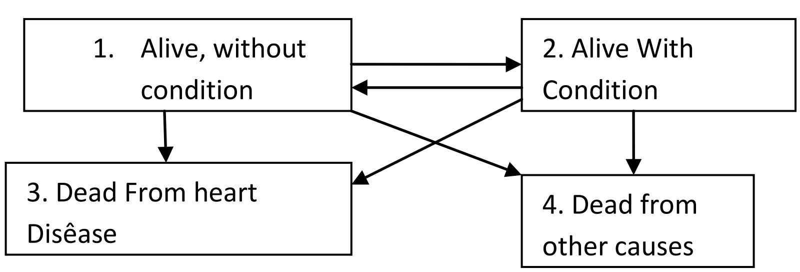
**(ii)**
In the case where the probabilities $P\ (X_t = j \mid X_s = i\ )$ for $i$, $j$ in S and 0 ≤$s$< $t$ depend only on the length of time interval $t$ $s$ , the process is called time-homogeneous. **[2]**

**(iii)** A model with time-inhomogeneous rates has more parameters, and there may not be sufficient data available to estimate these parameters. Also, the solution to Kolmogorov equations may not be easy (or even possible) to find analytically. **[2]**

**[8 Marks]**

**Solution 10 :**

**(i)**




**[2]**

**(ii)** With the state numbers in the diagram above we can write:
EITHER
Using the Markov assumption

OR
The Chapman Kolmogorov equation is
${}_{dt+t}P_x^{24} = {}_tP_x^{21}\,{}_{dt}P_{x+t}^{14} + {}_tP_x^{22}\,{}_{dt}P_{x+t}^{24} + {}_tP_x^{23}\,{}_{dt}P_{x+t}^{34} + {}_tP_x^{24}\,{}_{dt}P_{x+t}^{44}$

But ${}_{dt}P_{x+t}^{34} = 0$
And ${}_{dt}P_{x+t}^{44} = 1$

So:

${}_{dt+t}P_x^{24} = {}_tP_x^{21}\,{}_{dt}P_{x+t}^{14} + {}_tP_x^{22}\,{}_{dt}P_{x+t}^{24} + {}_tP_x^{24}$

Assuming that for small dt

${}_{dt}P_{x+t}^{ij} = \mu_{x+t}^{ij}dt + o(dt)$

Where $\lim_{dt-0} o\,dt/dt = 0$

then substituting, we have
${}_{dt+t}P_x^{24} = {}_tP_x^{21}\,\mu_{x+t}^{14}dt + {}_tP_x^{22}\,\mu_{x+t}^{24}dt + {}_tP_x^{24} + O(dt)$
so that
${}_{dt+t}P_x^{24} - {}_tP_x^{24} = {}_tP_x^{21}\,\mu_{x+t}^{14}dt + {}_tP_x^{22}\,\mu_{x+t}^{24}dt + O(dt)$

and hence
$d/dt({}_tP_x^{24}) = \lim_{dt-0} ({}_{dt+t}P_x^{24} - {}_tP_x^{24}) / dt = {}_tP_x^{21}\,\mu_{x+t}^{14}dt + {}_tP_x^{22}\,\mu_{x+t}^{24}dt$ **[5]**

**(iii)**
The maximum likelihood estimate (MLE) of the death rate from heart disease for persons with the condition is 25 / 1139 = 0.02195

The MLE of the death rate from heart disease for persons without the condition is 10 / 2046 = 0.00489.

An estimate of the variance of the maximum likelihood estimator of the death rate from heart disease for persons with the condition is .00489 /2046 = .00000239

The null hypothesis *H*0 is that there is no difference between the means.

The variance of the difference between the two estimates is therefore:
0.000019271 + 0.00000239 = 0.000021659.

THEN EITHER

A 95% confidence interval around the difference is therefore:

$(0.02195 – 0.00489) \pm 1.96\sqrt{0.000021659}$
$= 0.01706 \pm 1.96*0.004654$
$= 0.01706 \pm 0.009122$
$= (0.007938, 0.026182)$
which does not include zero
OR
Under the null hypothesis the difference ~ Normal (0, 0.000021659).
Our observed value of the difference is 0.021949 - 0.004888 = 0.01706
A $z$-score for the actual difference of 0.01706 is therefore
$(0.01706/\sqrt{0.000021659}) = 3.67$
and since this is greater than 1.96 we reject the null hypothesis at the 95% level
THEN
so the difference is statistically significantly different from zero **[5]**

**[12 Marks]**

******************************************************