BHARATHIDASAN UNIVERSITY

Tiruchirappalli- 620024

Tamil Nadu, India.

Programme: M.Sc. Statistics

Course Title: Survival Analysis and Clinical Trials

Course Code: 23ST04DEC

Unit-III

Semi-parametric Survival Models

Dr. T. Jai Sankar
Associate Professor and Head
Department of Statistics

Ms. I. Angel Agnes Mary
Guest Faculty
Department of Statistics

UNIT – III

Semi-parametric Model

Introduction to Semi-parametric regression for failure rate

A parametric survival model is one in which survival time (the outcome) is assumed to follow a known distribution. Examples of distributions that are commonly used for survival time are: Weibull, exponential (a special case of the Weibull), Gamma, log-normal, etc.

The Cox proportional hazards model, by contrast, is not a fully parametric model. Rather it is a semi-parametric model because even if the regression parameters (the betas) are known, the distribution of the outcome remains unknown. The baseline survival (or hazard) function is not specified in a Cox model (we do not assume any shape or form).

Let T denote the time to some event. Our data, based on a sample of size n, consists of the triple ($\tilde{T}_i$, Δ_i, X_i, i=1,...,n where $\tilde{T}_i$ is the time on study for the i-th patient, Δ_i is the event indicator for the i-th patient (Δ_i=1 if the event has occurred and Δ_i=0 if the lifetime is right-censored) and $X_i=(X_{i1},\ldots,X_{ip})^t$ is the vector of covariates or risk factors for the i-th individual which may affect the survival distribution of T.

Note that the covariates X_{ij}, with j=1, 2, …, p, may be time-dependent as $X_i(t)=(X_{i1},\ldots,X_{ip})^t$ whose value changes over time. This situation must be analyzed using the Extended Cox PH model. However, for ease of presentation, we shall consider the fixed-covariate case.

The Cox PH regression model (Cox 1972) is usually written in terms of the hazard model formula as follows

$$h(t) = h_0(t) \times exp(b_1x_1 + b_2x_2 + \ldots + b_px_p)$$

where,

- *t* represents the survival time
- h(t) is the hazard function determined by a set of p covariates ($x_1,x_2,...,x_p$)
- the coefficients ($b_1,b_2,...,b_p$) measure the impact (i.e., the effect size) of covariates.
- the term h0 is called the baseline hazard. It corresponds to the value of the hazard if all the x_i are equal to zero (the quantity exp(0) equals 1). The 't' in h(t) reminds us that the hazard may vary over time.

This model gives an expression for the hazard at time t for an individual with a given specification of a set of explanatory variables denoted by the bold X.

Based on this model we can say that the hazard at time t is the product of two quantities:

- The first of these, $h_0(t)$, is called the baseline hazard function or the hazard for a reference individual with covariate values 0.
- The second quantity is a parametric component which is a linear function of a set of p explanatory X variables that is exponentiated (it will be the *relative risk* associated with covariate values X).

Note that an important feature of this model, which concerns the proportional hazards (PH) assumption, is that the baseline hazard is a function of t, but does not involve the covariates. By contrast, the exponential expression involves the X's but not the time. The covariates here have a multiplicative effect and are called time-independent.

Note that the model is assuming proportional hazards (the hazard for any individual i is a fixed proportion of the hazard for any other individual j), that is:

$$\frac{h_i(t|\mathbf{X}_i)}{h_j(t|\mathbf{X}_j)} = exp(\boldsymbol{\beta}(\mathbf{X}_i - \mathbf{X}_j))$$

or

$$h_i(t|\mathbf{X}_i) = \exp(\boldsymbol{\beta}(\mathbf{X}_i - \mathbf{X}_j))h_j(t|\mathbf{X}_j)$$

So hazard functions for each individual should be strictly parallel and the hazard ratio is constant over time.

Cox's proportional hazards (PH) model with one and several covariates

The **Cox proportional-hazards model** (Cox, 1972) is essentially a regression model commonly used statistical in medical research for investigating the association between the survival time of patients and one or more predictor variables.

The purpose of the model is to evaluate simultaneously the effect of several factors on survival. In other words, it allows us to examine how specified factors influence the rate of a particular event happening (e.g., infection, death) at a particular point in time. This rate is

commonly referred as the hazard rate. Predictor variables (or factors) are usually termed *covariates* in the survival-analysis literature.

The Cox model can be written as a multiple linear regression of the logarithm of the hazard on the variables x_i, with the baseline hazard being an 'intercept' term that varies with time.

The quantities $\exp(b_i)$ are called hazard ratios (HR). A value of b_i greater than zero, or equivalently a hazard ratio greater than one, indicates that as the value of the i-th covariate increases, the event hazard increases and thus the length of survival decreases.

Put another way, a hazard ratio above 1 indicates a covariate that is positively associated with the event probability, and thus negatively associated with the length of survival.

In summary,

- HR = 1: No effect
- HR < 1: Reduction in the hazard
- HR > 1: Increase in Hazard

Note that in cancer studies:

- A covariate with hazard ratio > 1 (i.e.: b > 0) is called bad prognostic factor
- A covariate with hazard ratio < 1 (i.e.: b < 0) is called good prognostic factor

A key assumption of the Cox model is that the hazard curves for the groups of observations (or patients) should be proportional and cannot cross.

Consider two patients k and k' that differ in their x-values. The corresponding hazard function can be simply written as follow:

- Hazard function for the patient k:

$$h_k(t) = h_0(t)e^{\sum_{i=1}^{n} \beta x}$$

- Hazard function for the patient k':

$$h_{k'}(t) = h_0(t)e^{\sum_{i=1}^{n} \beta x'}$$

- The hazard ratio for these two patients $\left[\frac{h_k(t)}{h_{k'}(t)} = \frac{h_0(t)e^{\sum_{i=1}^{n} \beta x}}{h_0(t)e^{\sum_{i=1}^{n} \beta x'}} = \frac{e^{\sum_{i=1}^{n} \beta x}}{e^{\sum_{i=1}^{n} \beta x'}}\right]$ is independent of time t.

Variables selection for the Cox proportional hazard model

It is possible to improve the Cox proportional hazards model by selecting the variables being part of the model. There are two options to select the variables:

- Forward selection: The selection process starts by adding the variable with the largest contribution to the model. If a second variable is such that its entry probability is greater than the entry threshold value, then it is added to the model. This process is iterated until no new variable can be entered in the model.
- Backward selection: This method is similar to the previous one but starts from a complete model.

Example

The remission times (weeks) for two groups of leukemia patients are given below:

Group1	t (weeks)	6	6	6	7	10	13	16	22	23	6+	9+
	Log WBC	2.31	4.06	3.28	4.43	2.96	2.88	3.6	2.32	2.57	3.2	2.8
Group2	t (weeks)	1	1	2	2	3	4	4	5	5	8	8
	Log WBC	2.8	5	4.91	4.48	4.01	4.36	2.42	3.49	3.97	3.52	3.05

Group1	t (weeks)	10+	11+	17+	19+	20+	25+	32+	32+	34+	35+
	Log WBC	2.7	2.6	2.16	2.05	2.01	1.78	2.2	2.53	1.47	1.45
Group2	t (weeks)	8	8	11	11	12	12	15	17	22	23
	Log WBC	2.32	3.26	3.49	2.12	1.5	3.06	2.3	2.95	2.73	1.97

To perform the Cox-Proportional Hazard model.

Calculation

Let us consider two explanatory variables as predictors of survival time T, where T denotes "weeks until going out of remission." We label the explanatory variables X_1 (for group status) and X_2 (for log WBC) and consider a third variable that is, the product of X_1 and X_2.

Fitting three different Cox proportional hazards models are calculated,

Model 1:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	1.509	0.410	3.68	0.000	4.523	2.027	10.094
No. of subjects = 42		Log likelihood = −86.380			Prob > chi2 = 0.0001		

Model 2:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	1.294	0.422	3.07	0.002	3.648	1.595	8.343
log WBC	1.604	0.329	4.87	0.000	4.975	2.609	9.486
No. of subjects = 42		Log likelihood = −72.280			Prob > chi2 = 0.0000		

Model 3:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	2.355	1.681	1.40	0.161	10.537	0.391	284.201
log WBC	1.803	0.447	4.04	0.000	6.067	2.528	14.561
Rx x log WBC	−0.342	0.520	−0.66	0.510	0.710	0.256	1.967
No. of subjects = 42		Log likelihood = −72.066			Prob > chi2 = 0.0000		

$e^{Coef.}$ gives an estimated hazard ratio (HR) = exp(1.294) = 3.647

Z Wald statistic = Coef. / Std.Err. = 1.294/0.422 = 3.07

LR statistic: Log likelihood = -72.280

= -2 ln L(**log likelihood statistic**)

= -2 × (-72.280) = 144.56

In general:

$LR = -2 \ln L_R - (-2 \ln L_F)$

LR (**interaction in model 3**) = $-2 \ln L_{model\ 2} - (-2 \ln L_{model\ 3})$

= (-2 – (-72.280)) - (-2 – (-72.066))

= 144.550 - 144.132

= 0.428

LR is χ^2 with 1 d.f. under H0: no interaction.

0.40 < P < 0.50, not significant

Wald test P = 0.510

Test for treatment effect:

Wald statistic: P = 0.002 (highly significant)

LR statistic: compare -2 log L from model 2 with -2 log L from model without Rx

Conclusion

Treatment effect is significant, after adjusting for log WBC.

Stratified Cox-Model

The "stratified Cox model" is a modification of the Cox proportional hazards (PH) model that allows for control by "stratification" of a predictor that does not satisfy the PH assumption. Predictors that are assumed to satisfy the PH assumption are included in the model, whereas the predictor being stratified is not included. We first consider stratifying on a single predictor and then later consider stratifying on two or more predictors. Further, we distinguish between the use of a "no-interaction" version of the stratified Cox model and an alternative approach that allows interaction.

General Stratified Cox-Model

$Z_1, Z_2, \ldots, Z_k$ be the set of k variables that does not satisfying the PH assumption and $X_1, X_2, \ldots, X_p$ be the p variables satisfying the PH assumption.

To perform the stratified Cox procedure, we define a single new variable, which we call Z* which has k* categories where k* total number of combinations (strata), formed after categorising each of $Z_1, Z_2, \ldots, Z_k$.

For example, suppose k is 2, and Z_1 = age (an interval variable) and Z_2 = treatment status (a binary variable). Then we categorize age into, say, three age groups namely young, middle, and old and treatment status is binary variable with categories placebo and treatment.

		Age		
		Young	**Middle**	**Old**
Treatment group	**Treatment**	Strata 1	Strata 2	Strata 3
	Placebo	Strata 4	Strata 5	Strata 6

The general Stratified Cox model is given by

$$h_g(t, X) = h_{0g}(t) \times \exp[\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p]$$

Where, g = 1, 2,. . ., k*, strata defined by Z*.

Z* not included in the model

The baseline hazard function differ across strata i.e., $h_{0g}(t)$, g = 1, 2, . . . , k*, but $\beta_1, \beta_2, \ldots, \beta_p$ are fixed quantities for all the strata.

Since $h_{0g}(t)$ differs for each strata we would get different survival curves for each strata.

$$\text{Different baselines} \left\{ \begin{array}{lcl} \hat{h}_{01}(t) & \Rightarrow & \hat{S}_1(t) \\ \hat{h}_{02}(t) & \Rightarrow & \hat{S}_2(t) \\ \vdots & & \\ \hat{h}_{0k}(t) & \Rightarrow & \hat{S}_k(t) \end{array} \right\} \text{Different survival curves}$$

Note that the variable Z* (strata's of time dependent variable) is not explicitly included in the model, the baseline hazard function $h_{0g}(t)$ is allowed to be different for each stratum. And fitted Stratified Cox model will yield different estimated survival curves for each stratum because the baseline hazard functions are different for each stratum.

Stratified Cox model with no interaction

The stratified Cox model with no interactions is given by,

$$h_g(t, X) = h_{0g}(t) \times \exp[\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p]$$

Where, β coefficients do not vary over strata (no-interaction assumption).

In the above model we assume that there is no interaction effect between X_i and Z_i.

In other words, for a given $X_1, X_2, \ldots, X_p$ and $\exp[\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p]$ does not change across different strata defined by Z.

Stratified Cox model with interaction

The stratified Cox model with interactions is given by,

$$h_g(t, X) = h_{0g}(t) \times \exp[\beta_{1g} X_1 + \beta_{2g} X_2 + \ldots + \beta_{pg} X_p]$$

In the above model we consider that the effect of X_i on hazard at time t changes for the g^{th} stratum. In other words the effect of X_i changes as the strata changes which is the interaction between X_i and Z.

Alternative Stratified Cox model with interaction is modelled using the alternative approach which is given a follows:

If Z* takes k categories then defined k-1 dummy variables $Z_1, Z_2, \ldots, Z_{k-1}$.

Use the product terms involving X_i and Z_i. i.e., $X_i \times Z_i$

$$\begin{aligned} h_g(t,X) = h_{0g}(t) \times \exp\ [&\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \beta_{11}(Z_1^* \times X_1) + \beta_{21}(Z_1^* \times X_2) + \ldots + \\ &\beta_{p1}(Z_1^* \times X_p) + \beta_{12}(Z_2^* \times X_1) + \beta_{22}(Z_2^* \times X_2) + \ldots + \beta_{p2}(Z_2^* \times X_p) + \ldots + \\ &\beta_{1,k^*-1}(Z_{k^*-1}^* \times X_1) + \beta_{2,k^*-1}(Z_{k^*-1}^* \times X_2) + \ldots + \beta_{p1}(Z_{k^*-1}^* \times X_p) \end{aligned}$$

Testing for interaction effect

$$LR = -2 \ln L_R - (-2 \ln L_F)$$

The likelihood ratio (LR) test statistic is of the form -2 ln L_R minus -2 ln L_F, where R denotes the reduced model, which in this case is the no interaction model, and F denotes the full model, which is the interaction model.

$$H_0 : \begin{cases} \beta_{11} = \cdots = \beta_{p1} = 0 \\ \beta_{12} = \cdots = \beta_{p2} = 0 \\ \vdots \\ \beta_{1,k^*-1} = \cdots = \beta_{p,k^*-1} = 0 \end{cases} \qquad LR \sim \chi^2_{p(k^*-1)\ \mathrm{df}} \quad \text{under } H_0 : \text{no interaction}$$

Introduction to Competing risks analysis

Competing risks occur when there are at least two possible ways that a person can fail, but only one such failure type can actually occur.

For example,

1. Dying from either lung cancer or stroke
2. Advanced cancer patients either dying from surgery or getting hospital infection
3. Soldiers dying in accident or in combat
4. Limb sarcoma patients developing local recurrence, lung metastasis, or other metastasis over follow-up

A logical objective for competing risks data is to assess the relationship of relevant predictors to the failure rate or corresponding survival probability of any one of the possible events allowing for the competing risks of the other ways to fail. We might also want to compare the failure rates (e.g., using a hazard ratio) for two or more possible events, controlling for relevant predictors.

In the lung cancer versus stroke example above, we might ask whether the lung cancer death rate in "exposed" persons is different from the lung cancer rate in "unexposed" persons, allowing for the possibility that subjects could have died from a stroke instead.

We might also want to know if the lung cancer death rate differs from the stroke death rate controlling for predictors of interest.

In standard survival analysis these questions can be answered by using Kaplan Meier product limit method to obtain event probability over time, and Cox proportional hazard model to predict such probability. Likewise, in competing event data, the typical approach involves the use of KM estimator to separately estimate probability for each type of event, while treating the other competing events as censored in addition to those who are censored from loss to follow-up or withdrawal. This method of estimating event probability is called cause-specific hazard function, which is mathematically expressed as:

$$h_c(t) = \lim_{\Delta t \to 0} \frac{P(t \leqslant T_c < t + \Delta t | T_c \geqslant t)}{\Delta t}$$

The random variable T_c denotes the time to failure from event type c, therefore the cause-specific hazard function $h_c(t)$ gives the instantaneous failure rate at time t from event type c, given not failing from event c by time t.

Correspondingly, there is a cause-specific hazard model based on the Cox proportional hazard model which has the form of:

$$h_c(t, X) = h_{0c}(t) exp(\sum_{i=1}^{p} \beta_{ic} X_i)$$

This proportional hazard model of event type c at time t allows effects of the covariates to differ by event types, as the subscripted beta coefficient suggests.

Using these methods, one can separately estimate failure rate for each one of competing events. For instance, in our breast cancer mortality example, when death from breast cancer is the event of interest, the death from heart attack and all other causes should be treated as censored in addition to conventional censored observations. This would allow us to estimate the cause-specific hazard for breast cancer mortality rate, and go on to fit a cause-specific hazard model on breast cancer mortality. The same procedure can apply to death from heart attack when it becomes event of interest.

A major caveat of the cause-specific approach is that it still assumes independent censoring for subjects who are not actually censored but failed from competing events, as for standard censorship such as loss to follow up. Suppose this assumption is true, when focusing on cause-specific death rate from breast cancer, then any censored subject at time t would have the same death rate from breast cancer, regardless of whether the reason for censoring is either CVD or other cause of death, or loss to follow-up. This assumption is equivalent to saying competing events are independent, which is the foundation for the KM type of analysis to be valid. However, there is no way to explicitly test whether this assumption is satisfied for any given dataset. For instance, we can never determine whether a subject who died from heart attack would have died from breast cancer if he did not die from heart attack, since the possible death from cancer is unobservable for subjects died from heart attack. Therefore, estimates from cause-specific hazard function do not have an informative interpretation since it relies heavily on the independence censoring assumption.

The most popular alternative approach to analyze competing event data is called the Cumulative Incidence Function (CIF), which estimates the marginal probability for each competing event.

Cumulative Incidence Function (CIF)

The construction of a CIF is as straight forward as the KM estimate. It is a product of two estimates:

1) The estimate of hazard at ordered failure time tf for event-type of interest, expressed as:

$$\hat{h}_c(t_f) = m_{cf}/n_f$$

where the m_{cf} denotes the number of events for risk c at time t_f and n_f is the number of subjects at that time.

2) The estimate of overall probability of surviving previous time (t_{d-1}):

$$\hat{S}(t_{f-1})$$

where S(t) denotes the overall survival function rather than the cause specific survival function.

With these two estimates, we can compute the estimated incidence probability of failing from event-type c at time t_f as:

$$\hat{I}_c(t_f) = \hat{S}(t_{f-1}) \times \hat{h}_c(t_f)$$

The equation is self-explanatory: the probability of failing from event type c at time t_f is simply the product of surviving the previous time periods and the cause specific hazard at time t_f.

The CIF for event type c at time t_f is then the cumulative sum up to time tf (i.e., from f'=1 to f'=f) of these incidence probabilities over all event type c failure times, which is expressed as:

$$\mathbf{CIF}_c(t_f) = \sum_{f'=1}^{f} \hat{I}_c(t_f) = \sum_{f'=1}^{f} \hat{S}(t_{f'-1}) \times \hat{h}_c(t_{f'})$$

As we mentioned before, the CIF is equivalent to 1-KM estimator when there is no competing event. When there is competing event, the CIF differs from 1-KM estimator in that it uses overall survival function S(t) that counts failures from competing events in addition to the event of interest, whereas the 1-KM estimator uses the event-type specific survival function $S_c(t)$, which treats failures from competing events as censored.

By using the overall survival function, CIF bypasses the need to make unverifiable assumptions of independence of censoring on competing events. Since the S(t) is always less than $S_c(t)$, in competing event data, the CIF is always smaller than 1-KM estimates, which means the 1-KM tends to overestimate the probability of failure from the event type of interest.

Estimation problems in competing risk model for parametric model

Fine and Gray (1999), proposed a proportional hazards model aims at modeling the CIF with covariates, by treating the CIF curve as a subdistribution function. The subdistribution function is analogous to the Cox proportional hazard model, except that it models a hazard function (as known as subdistribution hazard) derived from a CIF. The Fine and Gray subdistribution hazard function for event type c can be expressed as:

$$h_{c,CIF}(t) = \lim_{\Delta \to 0} \frac{Pr(t < T_c < t + \Delta t | Tc > t \cup T_{c'} \leqslant t, c' \neq c)}{\Delta t}$$

The above function estimates the hazard rate for event type c at time t based on the risk set that remains at time t after accounting for all previously occurring event types, which includes competing events.

The CIF based proportional hazard model is then defined as:

$$h_{c,CIF}(t) = h_{0c,CIF}(exp)[\sum_{i=1}^{P} \gamma_i X_i]$$

This model satisfied the proportional hazard assumption for the subpopulation hazard being modeled, which means the general hazard ratio formula is essentially the same as for the Cox model, except a minor cosmetic difference that the betas in the Cox model is replaced by gammas

in Fine and Gray's model. Consequently, we should interpret the gammas in a similar way as we do for the betas estimated from a Cox model, except that it estimates the effect of certain covariates in the presence of competing events. The Fine and Gray model can also be extended to allow for time-dependent covariates.

Estimation problems in competing risk model for non- parametric model

Gray (1988), proposed a non-parametric test to compare two or more CIFs. The test is analogous to the log-rank test comparing KM curves, using a modified Chi-squared test statistic. This test does not require the independent censoring assumption.

Problem

Compute the CIC curves for each hospital using the following tables:

Hospital A

t_f	n_f	m_f	$\hat{h}_{ca}(t_f)$	$\hat{S}(t_{f-1})$	$\hat{I}_{ca}(t_f)$	$CIC(t_f)$
0	100	0	0	-	-	-
2	40	0	0	1	0	0
4	40	20	-	-	-	-

Hospital B

t_f	n_f	m_f	$\hat{h}_{ca}(t_f)$	$\hat{S}(t_{f-1})$	$\hat{I}_{ca}(t_f)$	$CIC(t_f)$
0	100	0	0	-	-	-
1	80	0	0	1	0	0
2	80	15	-	-	-	-
3	65	0	-	-	-	-
4	25	5	-	-	-	-

Procedure

- To calculate the estimate of hazard at ordered failure time t_f for event-type of interest is

$$\hat{h}_c(t_f) = m_{cf}/n_f$$

- To estimate of overall probability of surviving previous time (t_{d-1}) is

$$\hat{S}(t_{f-1})$$

- To compute the estimated incidence probability of failing from event-type c at time t_f is

$$\hat{I}_c(t_f) = \hat{S}(t_{f-1}) \times \hat{h}_c(t_f)$$

- To calculate Cumulative Incidence Function,

$$\mathbf{CIF}_c(t_f) = \sum_{f'=1}^{f} \hat{I}_c(t_f) = \sum_{f'=1}^{f} \hat{S}(t_{f'-1}) \times \hat{h}_c(t_{f'})$$

Solution

Hospital A

The estimate of hazard at ordered failure time t_f for event-type of interest is

$$\hat{h}_c(t_f) = m_{cf}/n_f$$

$h^{\wedge}_{ca}(4) = 20/40 = 0.5$

The estimate of overall probability of surviving previous time (t_{d-1}) is

$$\hat{S}(t_{f-1})$$

$S^{\wedge}(4) = 40/100 = 0.4$

The estimated incidence probability of failing from event-type c at time t_f is

$$\hat{I}_c(t_f) = \hat{S}(t_{f-1}) \times \hat{h}_c(t_f)$$

$I^{\wedge}_{ca}(4) = 0.5 \times 0.4 = 0.20$

t_f	n_f	m_f	$h^{\wedge}ca(t_f)$	$S^{\wedge}(t_{f-1})$	$I^{\wedge}_{ca}(t_f)$	$CIC(t_f)$
0	100	0	0	-	-	-
2	40	0	0	1	0	0
4	40	20	**0.5**	**0.4**	**0.20**	**0.20**

Hospital B

The estimate of hazard at ordered failure time t_f for event-type of interest is

$$\hat{h}_c(t_f) = m_{cf}/n_f$$

$h^{\wedge}_{ca}(2) = 15/80 = 0.1875$

$h^{\wedge}_{ca}(3) = 0/65 = 0$

$h^{\wedge}_{ca}(2) = 5/25 = 0.20$

The estimate of overall probability of surviving previous time (t_{d-1}) is

$$\hat{S}(t_{f-1})$$

$S^{\wedge}(2) = 80/100 = 0.8$

$S^{\wedge}(3) = 65/100 = 0.65$

$S^{\wedge}(4) = 25/100 = 0.25$

The estimated incidence probability of failing from event-type c at time t_f is

$$\hat{I}_c(t_f) = \hat{S}(t_{f-1}) \times \hat{h}_c(t_f)$$

$I^{\wedge}_{ca}(2) = 0.1875 \times 0.8 = 0.15$

$I^{\wedge}_{ca}(3) = 0 \times 0.65 = 0$

$I^{\wedge}_{ca}(4) = 0.20 \times 0.25 = 0.05$

t_f	n_f	m_f	$h^{\wedge}_{ca}(t_f)$	$S^{\wedge}(t_{f-1})$	$I^{\wedge}_{ca}(t_f)$	$CIC(t_f)$
0	100	0	0	-	-	-
1	80	0	0	1	0	0
2	80	15	**0.1875**	**0.8**	**0.15**	**0.15**
3	65	0	**0**	**0.65**	**0**	**0.15**
4	25	5	**0.20**	**0.25**	**0.05**	**0.20**

Result:

The CIC curves for each hospital using the following tables:

Hospital A

t_f	n_f	m_f	$h^{\wedge}ca(t_f)$	$S^{\wedge}(t_{f-1})$	$I^{\wedge}_{ca}(t_f)$	$CIC(t_f)$
0	100	0	0	-	-	-
2	40	0	0	1	0	0
4	40	20	**0.5**	**0.4**	**0.20**	**0.20**

Hospital B

t_f	n_f	m_f	$h^{\wedge}_{ca}(t_f)$	$S^{\wedge}(t_{f-1})$	$I^{\wedge}_{ca}(t_f)$	$CIC(t_f)$
0	100	0	0	-	-	-
1	80	0	0	1	0	0
2	80	15	**0.1875**	**0.8**	**0.15**	**0.15**
3	65	0	**0**	**0.65**	**0**	**0.15**
4	25	5	**0.20**	**0.25**	**0.05**	**0.20**