Journal of Modern Applied Statistical Methods

Volume 9 | Issue 2 Article 15

11-1-2010

Incidence and Prevalence for A Triply Censored Data

Hilmi F. Kittani
The Hashemite University, Jordan, kittanih@hu.edu.jo

DOI: 10.22237/jmasm/1288584840

Journal of Modern Applied Statistical Methods
November 2010, Vol. 9, No. 2, 480-487

1538 – 9472/10/$95.00

Incidence and Prevalence for A Triply Censored Data

Hilmi F. Kittani
The Hashemite University,
Jordan

The model introduced for the natural history of a progressive disease has four disease states which are expressed as a joint distribution of three survival random variables. Covariates are included in the model using Cox's proportional hazards model with necessary assumptions needed. Effects of the covariates are estimated and tested. Formulas for incidence in the preclinical, clinical and death states are obtained, and prevalence formulas are obtained for the preclinical and clinical states. Estimates of the sojourn times in the preclinical and clinical states are obtained.

Key words: Progressive disease model, prevalence, incidence, trivariate hazard function, censored data, proportional hazards model, sojourn times, chronic habitué.

Introduction

Louis, et al. (1978) introduced a natural history model for a progressive disease in a set of three articles: Albert, Gertman and Louis (1978), Albert, Gertman, Louis and Liu(1978) and Louis, Albert and Heghinian (1978). This model was extended by Kittani (1995a). Clayton (1978) also developed a model for association for the bivariate case and Oakes (1982) made inferences about the association parameter in Clayton's model. Clayton and Cuzick (1985) introduced the bivariate survival function for two failure times and made inferences about the association parameter, γ. Kittani (1995a, 1996, 1997, 1997-1998) considered the model for the bivariate case – that is, a case with two failure times (X, T) – by including covariates and by using Cox's proportional hazards model.

The motivation for this research lies in the fact that it is necessary to identify a three dimensional survival function for three failure times (X, Y, D) with four disease states (disease free state, preclinical state, clinical state and death state). In the model, X is the age upon entering the preclinical state (tumor onset or first heart attack), T is the age when entering the clinical state (symptoms first appear or second heart attack) and D is the age upon entering the death state (dying of cancer or acute myocardial infarction). Kittani (2010) considered estimating the parameters using nonparametric approach for a triply censored data.

Hilmi Kittani is a Professor of Statistics in the College of Science, Department of Mathematics. Email: kittanih@hu.edu.jo.

Background and Assumptions

As in the Louis, et al. (1978) model, it is assumed that $f_{XYZ}(x,y,z,a)$ is continuous – that is, $X = Y = Z = \infty$ is not allowed – and Y and Z are termed the sojourn times in the preclinical and clinical states respectively. The model proposed by Louis, et al. (1978) makes the assumption of no cohort effect, meaning that the distribution of the random variables (X, Y, Z) is independent of the age distribution A, or

$$f_{XYZA}(x, y, z, a) = f_{XYZ}(x, y, z) \times f_A(a)$$

and

$$f_{XYA}(x, y, a) = f_{XY}(x, y) \times f_A(a)$$

where $f_{XYZ}(x, y, z)$ is the joint pdf of (X, Y, Z) , $f_{XY}(x,y)$ is the joint pdf of X, Y and $f_A(a)$ is the pdf of A(the age distribution of the subject population). In addition, a subject is a chronic habitué of the PCS if, for that subject, $X < \infty$, $Y = \infty$, for example, subject never leaves PCS. According to the model, there will be no chronic habitués of the PCS or CS because, if a subject

lives long enough, then he/she will progress to the next state eventually (Louis et al., 1978).

The X, Y and Z axes are partitioned into I, J and K intervals according to Chiang, et al. (1989) and Hollford (1976); they assumed constant baseline hazards in each subinterval, $\lambda_{1i}(x) = \mu_{1i}$, $x \in I_i$, $\lambda_{2j}(y) = \mu_{2j}$, $y \in I_j$ and $\lambda_{3k}(z) = \mu_{3k}$, $z \in I_k$ in the i^{th}, j^{th} and k^{th} intervals respectively. The hazard functions for the n^{th} individual whose (X, Y, Z) values fall in the cube $I_i x I_j x I_k$ are modeled by assuming Cox's (1972) proportional hazards model and holds for each X, Y and Z in each respective I_i, I_j and I_k interval.

Assuming $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\eta}$ (regression parameters) for the covariate $\boldsymbol{\omega}$ (p-dimensional) are constant (the same) for all intervals to be estimated**.** The hazard functions λ_1, λ_2 and λ_3 in the I^{th}, J^{th} and K^{th} intervals for n^{th} individual whose observed (X, Y, Z) value is (X_n, Y_n, Z_n) will be defined as

$$\lambda_{1i}(x_n) = \mu_{1i} e^{\alpha' \omega_n}, x_n \in I_i = (a_i, a_{i+1}], \lambda_{2j}(y_n) = \mu_{2j} e^{\beta' \omega_n},$$

and

$$y_n \in I_j = (b_j, b_{j+1}], \lambda_{3k}(z_n) = \mu_{3k} e^{\eta' \omega_n}, z_n \in I_k = (c_k, c_{k+1}].$$

Where μ_{1i}, μ_{2j} and μ_{3k} are baseline hazard functions associated with X, Y and Z respectively. Assuming α, β and η are constant (the same) regression parameters for the covariate $\boldsymbol{\omega}$ for all intervals and to be estimated along with the association parameter γ.

The joint survival function for the three non-negative random variables (X, Y, Z) given by Kittani (1995b) is:

$$F(x, y, z) = [e^{\gamma \Lambda_1(x)} + e^{\gamma \Lambda_2(y)} + e^{\gamma \Lambda_3(z)} - 2]^{-\frac{1}{\gamma}}. \quad (2.1)$$

Where $\gamma > 0$, $x > 0$, $y > 0$, $z > 0$, and Λ_1, Λ_2, Λ_3 are the cumulative hazard functions associated with X, Y and Z respectively. For example, to compute $\Lambda_{1i}(x)$, which is the cumulative hazard function for the n^{th} individual whose x value falls in the i^{th} interval (assuming a constant hazard over each interval) is as follows:

$$\Lambda_{1i}(x_n) = \int_0^{x_k} \lambda_1(u)\,du = \left[\sum_{r=1}^{i-1} \mu_{1r}(a_{r+1} - a_r) + \mu_{1i}(x_n - a_i) \right] e^{\alpha' \omega} \quad (2.2)$$

where $\Lambda_{2j}(y_n)$ and $\Lambda_{3k}(z_n)$ are defined in a similar way. Thus, the joint density function (X, Y, Z) is

$$f(x, y, z) = \left[(\gamma+1)(\gamma+2)\lambda_1(x)\lambda_2(y) \; \lambda_3(z)\, e^{\gamma[\Lambda_1(x) + \Lambda_2(y) + \Lambda_3(z)]}\, U^{(-\frac{1}{\gamma} - 3)} \right] \quad (2.3)$$

where $\gamma > 0$, $x > 0$, $y > 0$, $z > 0$, and λ_1, λ_2 and λ_3 are base line hazard functions associated with X, Y and Z respectively as

$$U = e^{\gamma \Lambda_1(x)} + e^{\gamma \Lambda_2(y)} + e^{\gamma \Lambda_3(z)} - 2.$$

Kittani (1996) derived the likelihood function for the uncensored and censored cases in order to estimate the regression parameters by maximizing the likelihood function, that is, the n^{th} individual that generates data vector $\underline{\mathbf{w}}_n$, and $L(\underline{\mathbf{w}}_n)$ is the likelihood function contribution for the n^{th} individual as:

$$L(w_1, w_2, \ldots, w_N) = \prod_{n=1}^{N} L(w_n).$$

This likelihood function is maximized with respect to the unknown parameter vector; $\underline{\theta} = (\gamma, \underline{\alpha}, \underline{\beta}, \underline{\eta}, \underline{\mu}_1, \underline{\mu}_2, \underline{\mu}_3)$ with dimension (3p + I + J+ K+1) where p is the number of covariates.

To apply the Kittani (1995b) formula in the likelihood function it is first modeled for (X, Y, Z), then the transformations X = X, T = X + Y and D = X + Y + Z are performed to obtain the joint density function g(x, t, d) of (X, T, D) as:

$$g(x,t,d) = f(x,t-x,d-t) = (\gamma+1)(\gamma+2)\lambda_1(x)\lambda_2(t-x)\lambda_3(d-t).$$

$$e^{\gamma\left[\begin{matrix}\Lambda_1(x)+\Lambda_2(t-x)\\ +\Lambda_3(d-t)\end{matrix}\right]}\left[\begin{matrix}[e^{\gamma\Lambda_1(x)}+e^{\gamma\Lambda_2(t-x)}\\ +e^{\gamma\Lambda_3(d-t)}-2]^{(-\frac{1}{\gamma}-3)}\end{matrix}\right] \tag{2.4}$$

Preclinical, Clinical and Death Incidence

According to Louis, et al. (1978) under the assumption of no cohort effect, that is, (X, Y, Z) is independent of A, assuming $I_{PC}(a) = f_X(a)$, then preclinical incidence among those aged A is defined in terms of this model as:

$$I_{PC}(a) = f_X(a) = \sum_{i=1}^{I} \mu_{1i}\, e^{\alpha'\omega} e^{-\Lambda_{1i}(a)}, \quad a \in I_i \tag{3.1}$$

where f_X is the marginal density of X. In order to define the overall preclinical incidence, I_{PC} in terms of this model, the distribution of A must be defined. It is assumed throughout this article that if A is uniformly distributed over an interval I as

$$f_A(a) = 1/Id_i, \quad a \in I_i \tag{3.2}$$

where I is the number of intervals on the x-axis and d_i is the length of interval i, then the overall preclinical incidence in terms of this model is

$$I_{PC} = \int_0^\infty f_X(x).f_A(x)\,dx = \sum_{i=1}^{I} \frac{1}{I d_i} \int_{I_i} \mu_{1i}.e^{\alpha'\omega}.e^{-\Lambda_{1i}(x)}\,dx = \sum_{i=1}^{I} \frac{1}{I d_i}\left(e^{-\Lambda_{1i}(a_i)} - e^{-\Lambda_{1i}(a_{i+1})}\right). \tag{3.3}$$

Similarly, if there is no cohort effect, then the clinical incidence among those aged A is defined in terms of this model as

$$I_{CL}(a) = f_T(a) = \int_0^a f(x, a-x)dx = \int_0^a (\gamma+1)\mu_{1i}\mu_{2j}\, e^{(\alpha'+\beta')\omega} e^{\gamma[\Lambda_{1i}(x)+\Lambda_{2j}(a-x)]} U^{(-\frac{1}{\gamma}-2)}\,dx$$

Where $U = e^{\gamma\Lambda_{1i}(x)} + e^{\gamma\Lambda_{2j}(a-x)} - 1$ and f_T is the marginal density of T = X + Y. This integral cannot be obtained in a closed form and should be evaluated numerically. The overall clinical incidence in terms of this model is

$$I_{CL} = \int_0^\infty f_T(t)\, f_A(t)dt = \int_0^\infty \left(\int_0^t f(x,t-x)dx\right) f_A(t)\,dt = \sum_{j=1}^{J}\left[\frac{1}{Jd_j}\int_{J_j}\left(\int_0^t f(x,t-x)dx\right)dt\right] \tag{3.5}$$

where J is the number of intervals on the y-axis, d_j is the length of interval j and

$$f(x,t-x) = \left[\begin{array}{l} (\gamma+1)\,\mu_{1i}\,\mu_{2j}\,e^{(\alpha'+\beta')}\omega\, e^{\gamma(\Lambda_{1i}(x)+\Lambda_{2j}(a-x))} \\ \left(e^{\gamma(\Lambda_{1i}(x)} + e^{\gamma\Lambda_{2j}(a-x)} - 1 \right)^{(-\frac{1}{\gamma}-2)} \end{array} \right]$$

The above integral cannot be obtained in a closed form and should be evaluated numerically. Equations (3.1) – (3.5) are similar to those given by Kittani (1997).

Similarly, if there is no cohort effect, then death incidence among those aged A is defined in terms of this model as

$$I_{DI}(a) = f_D(a) = \int_0^a \left(\int_0^t f(x,\ t-x,\ a-t)\,dx \right) dt \tag{3.6}$$

This integral cannot be obtained in closed form and should be evaluated numerically. The overall death incidence in terms of this model is

$$\begin{aligned} I_{DI} &= \int_0^\infty f_D(t)\, f_A(t)dt \\ &= \int_0^\infty \left(\int_0^a \left(\int_0^t f(x,t-x,a-t)\,dx \right) f_A(t)\,dt \right) da \\ &= \sum_{k=1}^{K} \left[\frac{1}{Kd_k} \int_{I_k} \int_0^a \left(\int_0^t f(x,t-x,a-t)\,dx \right) dt \right] \end{aligned} \tag{3.7}$$

where K is the number of intervals on the z-axis and d_k is the length of interval.

Preclinical and Clinical Prevalence

According to Louis, et al. (1978) under the assumption of no cohort effect, (X, Y, Z) is independent of A, and the assumption of no chronic habitués of the PCS, then preclinical prevalence among those aged a is

$$\Phi_{PC}(a) = \int_0^a \left(\int_{a-x}^\infty f_{XY}(x,y)dy \right) dx, \tag{4.1}$$

then preclinical prevalence among those aged a is defined in terms of this model as

$$\begin{aligned} &\Phi_{PC}(a) = \\ &\int_0^a \mu_{1i}\, e^{\alpha'}\omega e^{\gamma\Lambda_{1i}(x)} \\ &\left(e^{\gamma\Lambda_{1i}(x)} + e^{\gamma\Lambda_{2j}(a-x)} - 1 \right)^{(-\frac{1}{\gamma}-1)} dx. \end{aligned} \tag{4.2}$$

The integral cannot be obtained in a closed form and should be evaluated numerically.

The overall preclinical prevalence according to Louis, et al. (1978) is

$$\Phi_{PC} = \int_0^\infty \Phi_{PC}(a) f_A(a)\,da \tag{4.3}$$

and, under the assumption of no cohort effect and no chronic habitué s of the PCS, the overall preclinical prevalence is defined in terms of this model as

$$\begin{aligned} \Phi_{PC} &= \int_0^\infty \Phi_{PC}(a) f_A(a)\,da \\ &= \sum_{i=1}^{M} \left(\frac{1}{Md_i} \int_{J_i} \Phi_{PC}(a)da \right) \\ &= \sum_{i=1}^{M} \left[\frac{1}{Md_i} \right. \\ &\left. \int_{J_i} \left(\int_0^a \mu_{1i}\, e^{\alpha'}\omega e^{\gamma\Lambda_{1i}(x)} \left(e^{\gamma\Lambda_{1i}(x)} + e^{\gamma\Lambda_{2j}(a-x)} - 1 \right)^{(-\frac{1}{\gamma}-1)} dx \right) da \right] \end{aligned} \tag{4.4}$$

The integral cannot be obtained in a closed form and should be evaluated numerically. Thus, the

clinical prevalence among those aged a in terms of this model is

$$\Phi_{CS}(a)=\int_0^a\int_{a-x}^{\infty}\int_{d-a}^{\infty} f_X(x,y,z)\,dz\,dy\,dx \tag{4.5}$$

where f(x, y, z) is given by

$$f(x,y,z)$$

$$=(\gamma+1)(\gamma+2)\lambda_1(x)\lambda_2(y)\lambda_3(z)e^{\gamma[\Lambda_1(x)+\Lambda_2(y)+\Lambda_3(z)]}U^{(-\frac{1}{\gamma}-3)} \tag{4.6}$$

where

$$U=[e^{\gamma\Lambda_{1i}(x)}+e^{\gamma\Lambda_{2j}(y)}+e^{\gamma\Lambda_{3k}(z)}-2].$$

Therefore the overall clinical prevalence in terms of this model is

$$\begin{aligned}\Phi_{CS}&=\int_0^{\infty}\Phi_{CS}(a)f_A(a)\,da\\&=\int_0^{\infty}\int_0^{a}\int_{a-x}^{\infty}\int_{d-a}^{\infty} f_{XYZ}(x,y,z)\,dz\,dy\,dx\\&=\sum_{k=1}^{K}\left[\frac{1}{Kd_k}\int_{I_k}\left(\int_0^{a}\int_{a-x}^{\infty}\int_{d-a}^{\infty} f_{XYZ}(x,y,z)\,dz\,dy\,dx\,da\right)\right]\end{aligned} \tag{4.7}$$

where f(x,y,z) is given by equation (4.6); the integral cannot be obtained in a closed form and should be evaluated numerically.

Estimation of the Sojourn Times in the Preclinical and Clinical States

Louis, et al. (1978) defined the mean duration of a disease in the preclinical state as

$$E(Y|X<\infty)=\frac{\int_0^{\infty}\Phi_{PC}(a)da}{\int_0^{\infty}I_{PC}(a)da}. \tag{5.1}$$

However, according to this model, no cohort effect and no chronic habitué s of the PCS are assumed, thus, the quantity E[Y|X < ∞] will be E(Y) because P[X < ∞] is 1. Therefore, substituting for $I_{PC}(a)$ and $\Phi_{PC}(a)$ in the above formula results in

$$E(Y)=\frac{\sum_{j=1}^{N}\int_{I_j}\left(\int_0^{a}\mu_{1i}e^{\alpha'\omega}e^{\gamma\Lambda_{1i}(x)}U^{(-\frac{1}{\gamma}-1)}\right)da}{\sum_{i=1}^{M}\left(e^{\Lambda_{1i}(a_i)}-e^{\Lambda_{1i}(a_{i+1})}\right)} \tag{5.2}$$

This integral cannot be obtained in a closed form and should be evaluated numerically.

Defining the mean duration of the disease in the clinical state as

$$E(Z|Y<\infty)=\frac{\int_0^{\infty}\Phi_{CS}(a)da}{\int_0^{\infty}I_{CS}(a)da} \tag{5.3}$$

and, assuming no cohort effect and no chronic habitué s of the CS, the quantity E[Z| Y < ∞] will be E(Z), because P[Y < ∞] is 1. Thus, substituting for $I_{CS}(a)$ and $\Phi_{CS}(a)$ in the above formula results in

$$E(Z)=\frac{\int_0^{\infty}\left(\int_0^{a}\left[\int_{a-x}^{\infty}\left\{\int_{d-a}^{\infty}f_{XYZ}(x,y,z)\,dz\right\}dy\right]dx\right)da}{\int_0^{\infty}\left(\int_0^{a}(\gamma+1)\mu_{1i}\mu_{2j}e^{(\alpha'+\beta')\omega}e^{\gamma(\Lambda_{1i}(x)+\Lambda_{2j}(a-x))}U^{(-\frac{1}{\gamma}-2)}dx\right)da} \tag{5.4}$$

where f(x, y, z) is given by equation (4.6). The integrals cannot be obtained in a closed form and should be evaluated numerically.

Asymptotic Distributions of the Epidemiological Measures

In order to make inferences about the epidemiological measures obtained, it is necessary to find their distributions; the Delta Method (Bishop, et al., 1975) is applied to determine means and variances. The parameter vector to be estimated is $\underline{\boldsymbol{\theta}} = (\gamma, \underline{\boldsymbol{\alpha}}, \underline{\boldsymbol{\beta}}, \underline{\boldsymbol{\eta}}, \underline{\boldsymbol{\mu}}_1, \underline{\boldsymbol{\mu}}_2, \underline{\boldsymbol{\mu}}_3)$ with dimension (3p + M+ N+ K+1) where p is the number of covariates, dim($\underline{\boldsymbol{\mu}}_1$) = **M,** dim($\underline{\boldsymbol{\mu}}_2$) = **N,** dim($\underline{\boldsymbol{\mu}}_3$) = K and dim($\underline{\boldsymbol{\alpha}}$) = dim($\underline{\boldsymbol{\beta}}$) = dim($\underline{\boldsymbol{\eta}}$) = p. Because $\underline{\hat{\theta}}$ is the MLE for $\underline{\theta}$ and, from the properties of the MLE's, $\underline{\hat{\theta}}$ is approximately normal with mean $\underline{\theta}$ and the covariance matrix $I^{-1}[\ \underline{\theta}\ ]$ is the inverted covariance matrix of $\underline{\theta}$ obtained from maximizing the log likelihood function for the censored case.

If $g(\underline{\hat{\theta}})$ is any function of $\underline{\hat{\theta}}$, the approximate distribution of $g(\underline{\hat{\theta}})$ may be found by applying the Delta Method as

$$g(\hat{\theta}) \approx N\left(g(\theta), [\frac{\partial g(\theta)}{\partial \theta}]\,[I(\theta)]^{-1}\,[\frac{\partial g(\theta)}{\partial \theta}]' \right). \tag{6.1}$$

where

$$\frac{\partial g(\theta)}{\partial \theta} = \left(\frac{\partial g(\theta)}{\partial \theta_1}, \frac{\partial g(\theta)}{\partial \theta_2}, \ldots, \frac{\partial g(\theta)}{\partial \theta_\upsilon} \right),$$

$$dim(\theta) = \upsilon = 3p + M + N + K + 1$$

and the estimated variance of $g(\underline{\hat{\theta}})$ is

$$\left([\frac{\partial g(\theta)}{\partial \theta}][I(\theta)]^{-1}\,[\frac{\partial g(\theta)}{\partial \theta}]' \right)_{\theta=\hat{\theta}}.$$

As an example, the formulas for the derivatives of the preclinical incidence are derived as follows. The estimate of preclinical incidence among those aged a

$$g_\theta(a) = I_{PC}(a) = \sum_{i=1}^{M} \mu_{li} e^{\alpha'\omega} e^{\Lambda_{li}(a)}, \quad a \in I_i \tag{6.2}$$

$$g_{\hat{\theta}}(a) = \hat{I}_{PC}(a) = \sum_{i=1}^{M} \hat{\mu}_{li} e^{\hat{\alpha}'\omega} e^{\hat{\Lambda}_{li}(a)}, \quad a \in I_i \tag{6.3}$$

where

$$\hat{\Lambda}_{li}(a) = \left[\sum_{r=1}^{i-1} \hat{\mu}_{lr}(a_{r+1} - a_r) + \hat{\mu}_{li}(a - a_i) \right] e^{\hat{\alpha}'\omega}$$

Differentiating g with respect to $\underline{\boldsymbol{\theta}}$, results in

$$\frac{\partial g(\theta)}{\partial \gamma} = \frac{\partial g(\theta)}{\partial \beta_m} = \frac{\partial g(\theta)}{\partial \mu_{2r}} = 0, \tag{6.4}$$

and

$$\frac{\partial g}{\partial \alpha_m} = \sum_{i=1}^{M} \mu_{li}\, z\, e^{\alpha'\omega} e^{-\Lambda_{li}(a)} (1 - \Lambda_{li}(a)), \quad a \in I_i \tag{6.5}$$

$$\frac{\partial g}{\partial \mu_{ls}} = \begin{cases} \sum_{s=1}^{M} e^{\alpha'\omega} e^{-\Lambda_{ls}(a)} [1 - \mu_{ls}(a - a_s) e^{\alpha'\omega}], & s = i, a \in I_i \\ -\sum_{i=1}^{M} \mu_{li}\, e^{2\alpha'\omega} e^{-\Lambda_{li}(a)} (a_{s+1} - a_s), & s < i,\ a \in I_i \\ 0, & s > i,\ a \in I_i \end{cases} \tag{6.6}$$

To test the effect of the covariates on morbidity and mortality (getting into the PCS, CS and DS):

$$H_0 : \underline{\boldsymbol{\alpha}} = 0 \text{ vs. } H_1 : \underline{\boldsymbol{\alpha}} \neq 0,$$

$$H_0 : \underline{\boldsymbol{\beta}} = 0 \text{ vs. } H_1 : \underline{\boldsymbol{\beta}} \neq 0,$$

and

$$H_0 : \underline{\boldsymbol{\eta}} = 0 \text{ vs. } H_1 : \underline{\boldsymbol{\eta}} \neq 0. \tag{6.7}$$

then the standard errors of the estimates are obtained from $I^{-1}[\ \underline{\boldsymbol{\theta}}\ ]$ which is the inverted

Hessian matrix obtained by numerical integration from special software, such as IMSL routines. From the properties of the MLE estimates, under H_0: $\underline{\boldsymbol{\theta}}_i = 0$, $\hat{\theta}_i$ is approximately normal with mean zero and standard error SE($\underline{\boldsymbol{\theta}}_i$). The test statistic

$$Z_{\theta_i} = \frac{\theta_i}{\mathrm{SE}(\theta_i)}. \tag{6.8}$$

is used to test the previous hypotheses and confidence intervals for $\underline{\boldsymbol{\theta}}_i$ can be obtained.

The estimate of the overall preclinical incidence

$$g(\theta) = I_{PC} = \sum_{i=1}^{M} \frac{1}{M d_i} \left(e^{-\Lambda_{1i}(a_i)} - e^{-\Lambda_{1i}(a_{i+1})} \right) \tag{6.9}$$

is

$$g(\hat{\theta}) = \hat{I}_{PC} = \sum_{i=1}^{M} \frac{1}{M d_i} \left(e^{-\hat{\Lambda}_{1i}(a_i)} - e^{-\hat{\Lambda}_{1i}(a_{i+1})} \right) \tag{6.10}$$

Differentiating g with respect to $\underline{\boldsymbol{\theta}}$, results in

$$\frac{\partial g(\theta)}{\partial \gamma} = \frac{\partial g(\theta)}{\partial \beta_m} = \frac{\partial g(\theta)}{\partial \mu_{2r}} = 0, \tag{6.11}$$

and

$$\frac{\partial g}{\partial \alpha_m} = \sum_{i=1}^{M} \frac{1}{M d_i} z e^{\alpha' \omega} \begin{pmatrix} \Lambda_{1i}(a_{i+1}) e^{-\Lambda_{1i}(a_{i+1})} \\ -\Lambda_{1i}(a_i) e^{-\Lambda_{1i}(a_i)} \end{pmatrix} \tag{6.12}$$

$$\frac{\partial g}{\partial \mu_{ls}} = \begin{cases} \sum_{i=1}^{M} \frac{1}{M} e^{\alpha' \omega} e^{-\Lambda_{1i}(a_{i+1})} & s = i,\ a \in I_i \\ \sum_{i=1}^{M} \frac{1}{M} e^{\alpha' \omega} \left(e^{-\Lambda_{1i}(a_{i+1})} - e^{-\Lambda_{1i}(a_i)} \right), & s < i,\ a \in I_i \\ 0, & s > i,\ a \in I_i \end{cases} \tag{6.13}$$

The covariance matrix for I_{PC} is

$$\left[\frac{\partial g(\theta)}{\partial \theta}\right] [I(\theta)]^{-1} \left[\frac{\partial g(\theta)}{\partial \theta}\right]'$$

where $I^{-1}[\underline{\theta}]$ the inverted covariance matrix of $\underline{\theta}$ obtained from maximizing the log likelihood function for the censored case.

References

Albert, A., Gertman, P. M., & Louis, T. A. (1978). Screening for the early detection of cancer I. The temporal natural history of a progressive of the disease state. *Mathematical Biosciences*, *40*, 61-59.

Albert, A., Gertman, P. M., Louis, T. A., & Liu, S. (1978). Screening for the early detection of cancer II. The temporal natural history of a progressive of the disease state. *Mathematical Biosciences*, *40*, 61-109.

Bishop, M., Feinberg, S., & Holland, P. (1975). *Discrete multivariate analysis*. Cambridge, MA: MIT Press.

Chaing, Y. K., Hardy, R. J., Hawkins, C. M., & Kapadia, A. S. (1989). An illness-death process with time dependent covariates. *Biometrics*, *45*, 669-681.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, *65*, 141-151.

Clayton, D. G., & Cuzick, J. (1985). Multivariate generalization of proportional hazards model. *Journal of the Royal Statistical Society*, *Series A*, *44*, 82-117.

Cox, D. R. (1972). Regression model and life tables (with discussion). *Journal of the Royal Statistical Society*, *Series B*, *34*, 187-220.

Hollford, T. R. (1976). Life tables with concomitant information (with discussion). *Biometrics*, *23*, 587-598.

Kittani, H. F. (1995a). Likelihood function for a progressive disease model for a bivariate survival function. *Qatar University Science Journal*, *15*(*2*), 287- 290.

Kittani, H. F. (1995b). Trivariate hazard function for censored survival data. *Journal of Mathematical Sciences*, *6*(*2*), 67- 73.

Kittani, H. F. (1996). Likelihood function for a progressive disease model for a trivariate survival function with covariates. *Journal of Mathematical Sciences*, *7*(*1*), 45-51.

Kittani, H. F. (1997). Epidemiological measures for a progressive disease model with a bivariate survival function with covariates. *Journal of Mathematical Sciences*, *8*(*1*), 41-50.

Kittani, H. F. (1997-1998). Likelihood function for the censored case for a progressive disease model for a bivariate survival function. *Aligarth Journal of Statistics*, *17* & *18*, 72-79.

Kittani, H.F. (2010). Estimation of Transition Functions for an Illness Death process. International Journal of Applied Mathematics and Statistics., Vol. 18, S10, 41-48.

Louis, T. A., Albert, A., & Heghinian, S. (1978). Screening for the early detection of cancer II. The temporal natural history of a progressive of the disease state. *Mathematical Biosciences*, *40*, 111-144.

Oakes, D. (1982). A model for association in bivariate survival data. *Journal of the Royal Statistical Society*, *Series B*, *44*, 412-422.