Journal of Economics and Administrative Sciences
2024; 30(140), pp. 450-465
P-ISSN 2518-5764
E-ISSN 2227-703X

Journal of Economics and Administrative Sciences (JEAS)

Available online at http://jeasiq.uobaghdad.edu.iq
DOI: https://doi.org/10.33095/619t1m17

A Partial Maximum Likelihood Method to Estimate Cox Model for Competing Risk Data with Application

Ibrahim Khalil Hussein*
Department of Statistics,
College of Administration & Economics,
University of Baghdad, Iraq
ibrahim.khaleel2101m@coadec.uobaghdad.edu.iq
*Corresponding author

Entsar Arebe Fadam
Department of Statistics,
College of Administration & Economics,
University of Baghdad, Iraq
entsar.arebe@coadec.uobaghdad.edu.iq

Received:26/9/2023 Accepted:20/12/2023 Published Online First: 30 /4/ 2024

Abstract:

Many Research and clinical studies have addressed the occurrence of failure (death). As a result of other (external) factors, which may introduce additional risks that compete with the event under study. The resulting data refers to the complete data on Competing risks, which are affected by time differences. The exact times of occurrence of these risks are known, meaning the times of failure are observed for all observations with certainty. The impact of these risks on the hazard function is estimated based on the Cox proportional hazards model, which is estimated using the partial maximum likelihood method and numerical algorithms for parameter estimation. This includes the effect of variables on the Cox hazard function and the nonparametric part, estimated by assessing the effect of time on the hazard function using the Kaplan-Meier formula and calculating competing risks through the cumulative hazard function.
These methods were applied to experimental data through large-scale simulations of different sizes and parameters and for several arithmetic means and standard deviations models. Moreover, applied to real data from a sample of 80 individuals with breast cancer. Analyzing the simulation results and real data revealed that the Downhill algorithm outperforms the Newton-Raphson algorithm in terms of estimation accuracy and efficiency, based on the statistical criterion for comparison, the root mean square error. In addition, competitive risks explained the effect of common variables in increasing competitive risks beyond the hazard function of the Cox model of the Newton-Raphson estimator. While it is converging to the hazard function of the Cox model for the Downhill estimator .

Paper type: Research paper
Keywords: Hazard function, Complete data, Cox proportional hazards model, partial Maximum likelihood, Newton-Raphson algorithm, Downhill algorithm.

P-ISSN 2518-5764
E-ISSN 2227-703X

1. Introduction:

Reliability and its applications are important topics with significant impact and relevance in various aspects of life. This importance is reflected in the study of survival and hazard functions for devices, equipment, and medical conditions under surveillance.

Competing risk methods are commonly used in biomedical Research, particularly in cancer, where dealing with multiple possible outcomes is crucial. For example, cancer-related deaths may be the primary event of interest. However, deaths due to other causes (e.g., liver function, diabetes, kidney function, etc.) are typical examples of competing risks. These risks vary depending on the specific condition under study, and the data rely on continuous monitoring during the specified study period.

Reliability has taken on an important and prominent role in estimating the effect of shared (external) variables and time on the hazard and survival functions. These functions depend on competing risk variables, the time of event discovery under study, and the time of study end.

Complexities arise when multiple time-varying covariates re[present various types or causes of failure that influence the cumulative event function. The application of the Cox proportional hazards model,was first proposed in 1972 by the renowned statistician D.R. Cox for estimating the impact of these shared variables on the hazard function. The most important thing that makes this model commonly used in analysis is the possibility of benefiting from it to estimate only part of its mediators.

The primary objective of estimating parameters in any model is to select the best-fitting model that yields good parameter estimates. This is achieved by choosing the optimal method and formula for obtaining parameter estimates based on the comparison criterion used in the study. There are several methods for estimating parameters in the Cox proportional hazards model to assess the impact of variables and time on the hazard function, affecting the survival function for the cases under study. By estimating the parametric part of (β), represented by the effect of variables on the risk function, using the partial maximum likelihood method, and to obtain the parameter results, used numerical algorithms (Newton Raphson - Down Hill).

The nonparametric part is estimated by assessing the effect of time on the baseline hazard function. Subsequently, competing risks are estimated using the cumulative hazard function, which is applied to Real complete data .

1.1 Literature review:

Many studies have discussed the issue of competing risks and their impact on the event function, including: Jewell et al. (2003) conducted a study on estimating survival distributions for current state data in the context of competing risks and observations were presented regarding the nonparametric maximum likelihood estimate, which is the estimate from the age distribution in both natural and surgical menopause. Groeneboom et al. (2008) used the nonparametric estimation of current status data with competing risks; The main focus was on the maximum likelihood estimator (MLE) and the naive estimator. It has been proven that both estimators converge globally and locally at a rate of n^(1/3). Tang et al. (2013) applied the Cox proportional hazards model was applied to analyze early failure data in power cables; The Cox model analyzes a set of common variables simultaneously and identifies the variables that have significant effects on cable failure. Hudgens et al. (2014) Studied the Parametric estimation of the cumulative event function (CIF) for interval observation competing risk data, which is based on prior nonparametric estimation, as well as the simple probability estimator, which uses only part of the observed data. The simple estimator gives a separate estimate for the models for each cause, unlike the complete maximum likelihood, which fits all models at one time. Mao and Lin (2017) used a semi-parametric regression model to formulate the effects of shared variables on the cumulative event function that needed estimation. They employed a wide range of semi-parametric transformation models that extend to Fine and Gray models. Do and Yang (2017) Suggested several ways to analyze competing risk data. In case of loss of failure causes and

P-ISSN 2518-5764
E-ISSN 2227-703X

failure time, the (Klein–Andersen's) pseudo-value approach was applied; The proposed method was evaluated through comparison with full case analysis in several simulation settings. Thackham and Ma (2020) proposed to apply the Cox model to deal with non-proportional hazards by estimating the partial likelihood of time-varying covariates. The lack of precision in estimating regression coefficients in small samples was addressed by developing the maximum likelihood. To estimate the regression coefficients and the basic hazard, shown through simulation to have increased accuracy compared to partial likelihood (PL) estimates in small samples. Guerrero et al. (2023) conducted an individual patient data (IPD) meta-analysis to assess the effect of anticoagulation on all-cause mortality in patients with cirrhosis and portal vein thrombosis; previous meta-analyses demonstrated the safety and efficacy of anticoagulation in the recanalization of portal vein thrombosis in patients with cirrhosis Whether this benefit translates into improved survival is unknown.

The problem Research is the occurrence of additional risks that affect the risk function, taking into account the actual failure time; these risks also affect the increase in the cumulative risk function represented by Competing risks over the model's risk function.

The Objective of Research is to estimate the effect of covariates represented by Competing risks on the model's risk function for complete data and determine the best method based on the statistical criterion, root mean square error.

2. Material and Methods:

2.1 Cox Proportional Hazards Model:

It is one of the widely used models in survival analysis and hazard function estimation for Censored data in the experiment, and in the mainly time-dependent areas of study. It is an alternative to commonly used models such, as the linear and logistic regression models, which are unsuitable for use with Censored data. This model is one of the popular models that study Censored data due to the ease of dealing with the data (Tang et al, 2013).

The Cox model is considered a semi-parametric model because it consists of two parts. The first part is parametric, representing the exponential function to estimate the parameter, and the other part is nonparametric, representing the effect of time on the basic hazard function. It is possible to benefit from the Cox model by estimating only part of its mediators, which makes it widely used. The mathematical formula for the Cox proportional hazards model is typically expressed as follows (Thackham et al., 2020).

$$h(t,x) = h_o(t).exp(\beta^T X) \quad (1)$$

Where:

$h_o(t)$ The baseline hazard function represents, β Cox Model Parameters Vector, X Matrix of Independent Variables Expected to Impact the Hazard Function $h(t,x)$

If you have two independent variables, X1 and X2, the Hazard Ratio for the first variable, X1, to the second variable, X2, is calculated as follows (Scheike and Sun, 2007).

$$\frac{h(t,x_1)}{h(t,x_2)} = \frac{h_o(t).\exp(\beta x_1)}{h_o(t).\exp(\beta x_2)}$$

$$\frac{h(t,x_1)}{h(t,x_2)} = \exp[\beta(x_1 - x_2)]$$

After taking the natural logarithm, it becomes:

$$ln\left[\frac{h(t,x_1)}{h(t,x_2)}\right] = \beta(x_1 - x_2)$$

This means that β represents the increase in the natural logarithm of the first variable's hazard ratio to the second variable's hazard by one unit. $\frac{h(t,x_1)}{h(t,x_2)}$

2.1.1 Estimation Methods for the Cox Model:

2.1.1.1 Partial Maximum Likelihood Method:

In 1980, the scientist Cox proposed a method for estimating the parameters β, called the Partial Maximum Likelihood Method. It is a widely used method that depends on the order of events (death) as follows: $t_{(1),}t_{(2)},..,t_{(r)}$, The sum of the elements under danger (death) at t is F (ti) The formula that illustrates the Partial Likelihood Method for parameter estimation is as follow (Cox, 1975).

$$f(\beta) = \frac{h(t_{(i)}, x_k)}{\sum_{j\in F(t_{(i)})} h(t_{(i)}, x_j)} \qquad (2)$$

Where $h(t_{(i)}, x_k)$ represents the studied event function at time t(i), and $\sum_{j\in F(t_{(i)})} h(t_{(i)}, x_j)$ represents the function of the occurrence of the event of interest (death) for each element under the risk set $F_{(t(i))}$.

Therefore, the partial likelihood function under the assumption of the event of interest (death) occurring at time t(i) can be written as follows (Sinha et al., 2003).

$$L(\beta) = \prod_{i=1}^{r} \frac{\exp(\beta^T x_k)}{\sum_{j\in F(t_{(i)})} \exp(\beta^T x_i)} \qquad (3)$$

By taking the logarithm of the partial likelihood function:

$$lnl(\beta) = \sum_{i=1}^{r} \{x_i\beta^T - ln[\sum_{j\in F} \exp(x_i\beta^T)]\} \qquad (4)$$

The parameter estimates from the partial likelihood can be obtained using numerical methods, and two commonly used numerical methods are the Newton-Raphson algorithm and the Downhill algorithm (DH) for the function (4).

a) Newton-Raphson Algorithm:

The Newton-Raphson algorithm is efficient for finding roots of real-valued functions and solving linear equations. It is also commonly used to solve nonlinear equations that may be complex and cannot be easily solved using traditional methods. This algorithm was originally proposed by Isaac Newton and Joseph Raphson in 1960(Yalçınkaya et al., 2018).

Using the Newton-Raphson algorithm in this context is to obtain the best parameter estimates β based on initial estimates (Initial estimates are obtained from the partial maximum Likelihood method). using the newton-Raphson algorithm's iterative approach for the partial likelihood equation defined in formula (4), the NR algorithm can be expressed mathematically as follows (Akram and Ann , 2015).

$$\hat{\beta}^{(r+1)} = \hat{\beta}^{(r)} - \frac{F^{(r)}}{P^{(r)}} \qquad (5)$$

In the context of the Newton-Raphson algorithm:

- r represents the number of iterations.
- $\hat{\beta}^{(r+1)}$represents the new parameter estimates
- $\hat{\beta}^{(r)}$represents the parameter estimates obtained in the r.
- $F^{(r)}$represents the first derivative vector of the logarithm of the partial likelihood function in the r-th iteration.(Casella and Bachmann, 2021).

$$F = \left(\frac{\partial LnL(\beta)}{\partial\beta}\right) = \sum_{i=1}^{r} x_i - \frac{\sum x_j e^{x_j\beta}}{\sum_{j\in F} e^{x_j\beta}}$$

$P^{(r)}$The matrix that represents the second derivative of the logarithm of the partial likelihood function for the r

$$P = \left[\frac{\partial^2 LnL(\beta)}{\partial^2 \beta}\right] = -\sum_{j\in F}\left[\frac{(\sum_{j\in F} e^{x_j\beta})(\sum_{j\in F} Xj^2 e^{x_j\beta}) - (\sum_{j\in F} Xje^{x_j\beta})^2}{(\sum_{j\in F} e^{x_j\beta})^2}\right]$$

After calculating the absolute difference between the new and previous parameter estimates, the algorithm checks whether this absolute difference is smaller than a predefined threshold value c. If the absolute difference is smaller than c, the algorithm stops, and the optimal solution is printed.

$$\left|\beta^{(r+1)} - \beta^{(r)}\right| < c \qquad (6)$$

Where c is an extremely small constant value.

b) Downhill Algorithm:

This algorithm is one of the numerical algorithms proposed by the two scientists Nelder-Mead in 1965. Which works to find the optimal solution to the objective function, that is, to obtain the minimum limit of the objective function in complex functions, which depends on guessing several points of the objective function (Fajfar et al., 2017). Moreover, because it does not need derivatives, it is very popular in many fields of science and technology, especially chemistry and medicine. The Downhill Algorithm is based on a geometric shape with several geometric forms of n dimensions and n+1 points represented as $(Z_1, Z_2, \ldots, Z_{n+1})$. These points represent the order of values of the objective function at each test, as follows (Galántai , 2021).

$$f_{z_1} \le f_{z_2} \le \cdots \le f_{z_{n+1}} \qquad (7)$$

Where Z1 represents the best point, and Z_{n+1}represents the worst point. Here are the steps for parameter estimation:

1.Choose the partial likelihood function, represented by the objective function, which is the negative logarithm of Equation (4).

$$f_z = -lnL(z) \quad , \qquad where\ Z = (\beta)$$

2.Fixing the algorithm parameters. Reflection $\alpha = 1$, Expansion $\varepsilon = 2$, Contraction $\gamma = 0.5$, Contraction $\psi = -0.5$.

3.Create an initial solution matrix **S** with dimensions of ((n+1)*1).

$$S = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{n+1} \end{bmatrix} \qquad (8)$$

4.Estimate the objective function for each row of the matrix S and arrange the estimates from lowest to highest.

$$f_{z_1} \le f_{z_2} \le \cdots \le f_{z_{n+1}}$$

5.The reflection point (r) is calculated using the following formula:

$$Z_r = \bar{S} + \alpha(\bar{S} - Z_{n+1}) \qquad (9)$$

After calculating the reflection point, we compute the objective function (f_{z_r}). If:

$f_{z_1} < f_{z_r} < \cdots < f_{z_n}$ We set $Z_r = Z_{n+1}$and then proceed to step (9). Otherwise, we move on to the next step.

6.Calculate the expansion point (e) according to the following formula:

$$Z_e = \bar{S} + \varepsilon(Z_r - \bar{S}) \qquad (10)$$

Calculate the value of the objective function (f_{z_e}) after expansion. If:

$f_{z_e} < f_{z_r}$ Calculate the point of contraction (c) using the formula: $Z_e = Z_{n+1}$Proceed to step 9.

7.Calculate the contraction point (c) using the formula:

$$Z_c = \bar{S} + \gamma(Z_{n+1} - \bar{S}) \qquad (11)$$

Calculate the objective function f_{z_c}after the contraction. If:

$f_{z_c} < f_{z_n}$ Exactly, $Z_c = Z_{n+1}$and proceed to step 9. If not, you move on to the next step.

8. Calculates the contraction point (Sh) based on the formula mentioned.

$$Z_{sh} = Z_1 + \psi(Z_i - Z_1) \qquad (12)$$

9. Applied when the stopping condition is met in the previous steps, as indicated by the formula.

$$\left|\frac{\max(f)-\min(f)}{\max(f)}\right| < \varepsilon$$

If the stopping condition is met, the algorithm proceeds to print the solution; otherwise, it returns to Step 5. This iterative process continues until the desired convergence is achieved (Mehta, 2020).

2.1.1.2 Kaplan-Meier:

The Kaplan-Meier estimator is a non-parametric estimator used to estimate the survival function. This estimator is known for its ease of calculation and interpretation. In medical research, it is often used to measure the proportion of patients who survive for a certain period after treatment. In other fields, it can be used to estimate the length of time individuals remain unemployed after job loss. It was named after Edward L. Kaplan and Paul Meier in 1958 (Smith and Smith, 2003).

The visual representation of this function is typically called the Kaplan-Meier curve, which shows the survival probability on the Y-axis and time on the X-axis. Kaplan assumed that the event of interest occurred at a specific time, and the probability of survival for all observations was the same, regardless of when they entered the study. Observations Submissive Censored have the same probability of survival.

The following formula defines the Kaplan-Meier survival function (KM) (Andrade, 2023).

$$\hat{S}(t)=\prod_{i=1}^{k}\frac{n_i-d_i}{n_i} \qquad ,k=1,2,\ldots,r \qquad (13)$$

the definition

n_i represents the number of individuals who have survived until time t_i.

d_i represents the number of events (deaths) at time t_i.

The following formula gives the cumulative failure distribution function:

$$\hat{F}(t)=1-\hat{S}(t)$$

The probability density function is estimated using the non-parametric empirical method according to the following formula:

$$\hat{f}(t)=-\frac{\hat{S}(t_{i+1})-\hat{S}(t_i)}{t_{i+1}-t_i}$$

$$=\frac{1}{(t_{i+1}-t_i)(n+1)} \qquad for \qquad t_{i<t<t_{i+1}} \qquad (14)$$

The following formula gives the failure time hazard function (Teoh, 2008).

$$\hat{h}(t)=\frac{\hat{f}(t)}{\hat{S}(t)}=\frac{1}{(t_{i+1}-t_i)(n+1-d_i)} \qquad for \quad t_{i<t<t_{i+1}} \qquad (15)$$

2.2 Competing risk

Competing risks arise when there are multiple possible outcomes in clinical research during survival analysis. For example, cancer-related deaths may be of primary interest, but deaths due to other causes not related to cancer are typical examples of competing risks. In such cases, death can occur before the disease onset (Sildnes, 2015). Competing risks can be estimated using the Cumulative Incidence Function (CIF), which can be estimated using the Nelson-Aalen estimator through the following formula (Groeneboom et al., 2008).

$$\hat{H}(t)=\sum\frac{d_i}{r_i} \qquad (16)$$

Where d_i is the number of individuals who die during the period t_i, and r_i is the number of individuals at risk during t_i.

3. Discussion of Results :

3.1 Simulation:

Simulation is representing or mimicking the real world, obtaining a similar or analogous model or system without relying on the same model or system. Often, we encounter complex processes in the real world that are difficult to analyze logically, leading to the translation of statistical theories through simulation to obtain results (estimates) that mimic the real world. Samples of different sizes (small, medium, large) are taken to visually represents the studied process or the real-world scenario through simulation. Simulation is used to compare different estimation methods and determine which is best for estimation (Thackham and Ma, 2020).

3.1.1 Simulation involves the following steps:

Stage 1:

Generating common variable values $(X_1, X_2, X_3)$for the complete data in the first stage involves random generation based on a uniform distribution over the interval (0, 1) for the initial value. Afterwards, the data is generated using a normal distribution with arithmetic means and standard deviations that have been assumed, as described by the following formula:

$$f_{(x)} = \frac{x - \mu}{\sigma} \quad ,$$

Where μ represents the arithmetic mean, σ represents the standard deviation.

As for time t, it was generated in vector form t=1,2,…,n .

Stage 2:

This stage is one of the most crucial, where the initial assumed values are determined. The values are specified as follows:

- Assumption of the required sample size: n=(20, 50, 90, 120).
- Determination of repetitions for each experimental unit: 1000
- Determine initial parameter values $\hat{\beta}$ for the Newton-Raphson estimator: (0.242, 1.342, 0.842) and for the Downhill estimator: (1.1778, 0.7661, 2.1023). The arithmetic means and variances are shown in the following table.

Table 1: Displays the arithmetic means and standard deviations of the common variables

Model	Mean(t)	$\mu(X_1)$	$\mu(X_2)$	$\mu(X_3)$	Std(t)	$\sigma(X_1)$	$\sigma(X_2)$	$\sigma(X_3)$
I	1584.35	5.85	23.03	4.83	318.36	1.87	16.57	1.68
II	1742.79	6.44	25.33	5.31	350.20	2.05	18.23	1.85
III	1425.92	5.27	20.73	4.34	286.53	1.68	14.91	1.51

Where $\mu(t)$ represents an arithmetic mean of time close to the time mean of real data.

Stage 3:

The parameters were estimated using numerical algorithms (Newton-Raphson and Downhill simplex) for the partial likelihood method, which is the parametric part. The non-parametric part was estimated automatically for different sample sizes using Kaplan-Meier. This was done to obtain the hazard function for the Cox model as indicated in Equation (1) and to calculate competitive risk and the reliability function based on Equations (16) and (13), respectively.

Stage 4:

Comparison between the studied estimation methods and determining the best approach was carried out based on the statistical measure Root Mean Square Error (RMSE) for the Cox model, according to the following formula:

$$RMSE = \sqrt{MSEh_{(t,x)}} \qquad (17)$$

where

$$MSEh_{(t,x)} = \frac{\sum_{i=1}^{r} (\hat{\hat{h}}_t - \hat{h}_t)^2}{r} \qquad (18)$$

Where:

- r: represents the number of repetitions for each experiment.
- $\hat{h}_t$: is the estimated hazard function for the initial assumed values.
- $\hat{\hat{h}}_t$: is the estimated hazard function for the Cox model.

3.1.2 Simulation Experiment Results:

Based on the results obtained from the various methods, different sample sizes, and different parameters models, a comparison will be made to determine the best approach based on the Root Mean Square Error (RMSE). These results are presented in the following tables:

The first case: In this scenario, the assumed parameters $\hat{\beta}(NR) = (0.242\,, 1.342\,, 0.842)$are used with three different models of arithmetic means and standard deviations denoted by symbols (III)(II)(I) for various sample sizes shown in Table 2 .

Table 2: Shows the RMSE values for the parameters and the hazard function of the Cox model for the assumed parameters and three different models, along with various sample sizes. For comparison between estimation methods.

Models of the initial parameters	N	Methods	RMSE(β_1)	RMSE(β_2)	RMSE(β_3)	RMSEh$_{(t,x)}$	Best Method
I	20	NR	0.8451	0.6087	0.4761	0.3018	DH
		DH	0.7761	0.9014	0.7889	0.0027	
	50	NR	0.8037	0.6287	0.4574	0.1787	DH
		DH	0.8086	0.8564	0.8708	0.0015	
	90	NR	0.7633	0.6616	0.4512	0.6378	DH
		DH	0.8235	0.8672	0.7966	0.0041	
	120	NR	0.7879	0.6396	0.4531	0.1366	DH
		DH	0.8202	0.8540	0.7937	0.0010	
I I	20	NR	0.7774	0.6494	0.4524	2.1849	DH
		DH	0.7881	0.7617	0.8389	0.0063	
	50	NR	0.8388	0.6038	0.4676	5.2208	DH
		DH	0.8586	0.8389	0.8959	0.0042	
	90	NR	0.7921	0.6388	0.4558	2.1726	DH
		DH	0.8681	0.8007	0.8077	0.0024	
	120	NR	0.7426	0.6802	0.4506	1.3282	DH
		DH	0.7951	0.8092	0.8439	0.0033	
I I I	20	NR	0.8372	0.6135	0.4732	1.4089	DH
		DH	0.8279	0.8533	0.9114	0.0055	
	50	NR	0.7662	0.6595	0.4518	0.4023	DH
		DH	0.8266	0.7942	0.8779	0.0025	
	90	NR	0.7697	0.6546	0.4506	0.0768	DH
		DH	0.8535	0.8684	0.8630	0.0010	
	120	NR	0.7564	0.6665	0.4499	0.3569	DH
		DH	0.7709	0.7313	0.8106	0.0043	

From Table (2), we observe that in all the models mentioned in Table (1), the vast majority of the estimates and risks are for the Cox model. The larger the sample size, the lower the estimator (RMSE), is consistent with statistical theory. Our results show that the Downhill algorithm outperforms the Newton-Raphson algorithm estimation accuracy and efficiency.

The second case: represents the assumed parameters $\hat{\beta}(DH) = (1.1778\,, 0.7661\,, 2.1023)$ with the three models for computational environments, standard deviations, and identified by symbols (III)(II)(I) for different sample sizes shown in Table 3 .

Table 3: Presents the (RMSE) values for the parameters and the hazard function for the Cox model with the assumed parameters and the three models of computational environments for various sample sizes for comparison between estimation methods.

Models of the initial parameters	N	Methods	RMSE(β_1)	RMSE(β_2)	RMSE(β_3)	RMSEh$_{(t,x)}$	Best Method
I	20	NR	0.6139	0.8685	0.8744	0.4925	DH
		DH	1.3932	1.3239	1.3425	0.0074	
	50	NR	0.6209	0.8813	0.8617	0.3031	DH
		DH	1.3660	1.3120	1.3349	0.0026	
	90	NR	0.6408	0.9112	0.8388	0.2420	DH
		DH	1.4150	1.3482	1.3691	0.0009	
	120	NR	0.6313	0.8978	0.8478	0.2724	DH
		DH	1.3849	1.3824	1.3599	0.0007	
I I	20	NR	0.6273	0.8897	0.8574	1.5389	DH
		DH	1.3714	1.3853	1.4064	0.0023	
	50	NR	0.6533	0.9347	0.8118	0.4272	DH
		DH	1.4214	1.3307	1.2541	0.0048	
	90	NR	0.6468	0.9233	0.8239	3.6767	DH
		DH	1.3921	1.3319	1.3706	0.0257	
	120	NR	0.6128	0.8653	0.8791	0.3605	DH
		DH	1.2863	1.4036	1.3936	0.0009	
I I I	20	NR	0.6979	0.9909	0.7818	0.8350	DH
		DH	1.3328	1.3799	1.4413	0.0097	
	50	NR	0.6627	0.9393	0.8241	0.2751	DH
		DH	1.3370	1.3888	1.3421	0.0023	
	90	NR	0.6447	0.9219	0.8221	0.2112	DH
		DH	1.3692	1.3665	1.3383	0.0002	
	120	NR	0.6892	0.9869	0.7680	0.3162	DH
		DH	1.3424	1.3589	1.2639	0.0008	

From Table (3), we observe that in all the models mentioned in Table (1), the majority of estimations and the hazard for the Cox model, the larger the sample size, the lower the (RMSE) estimator, which is consistent with statistical theory. The results show that the Downhill algorithm is the best, while the Newton-Raphson algorithm performed the worst in terms of estimation accuracy and efficiency.

3.2 Application of Real Data:

Data was collected from the Ministry of Health / Al-Amal Cancer Hospital for a sample of size n=80 individuals with breast cancer. Three variables (X_1: Blood Sugar, X_2: Liver Functions, X_3: Kidney Functions) represent competing risks, and the specified study period was from 2019 to 2023. Initial values for the parameters ($\beta_1 = 0.4$, $\beta_2 = 1.5, \beta_3 = 1$) were determined to obtain the initial estimator for the Cox model's hazard function. Afterwards, estimation methods were applied to obtain parameter estimators and the baseline hazard function for the time effect. The best method among the estimation techniques was determined based on the comparison criterion (RMSE) for the Cox model with formula (1) and is presented in the table (4) below:

Table 4: Presents the estimated parameters and RMSE for the Cox model's hazard function:

Methods	Beta estimated			$RMSEh_{(t,x)}$
	β_1	β_2	β_3	
Newton Raphson	0.2360	1.3360	0.8360	6.2085e-06
Downhill	1.1778	0.7661	2.1023	2.4107e-06

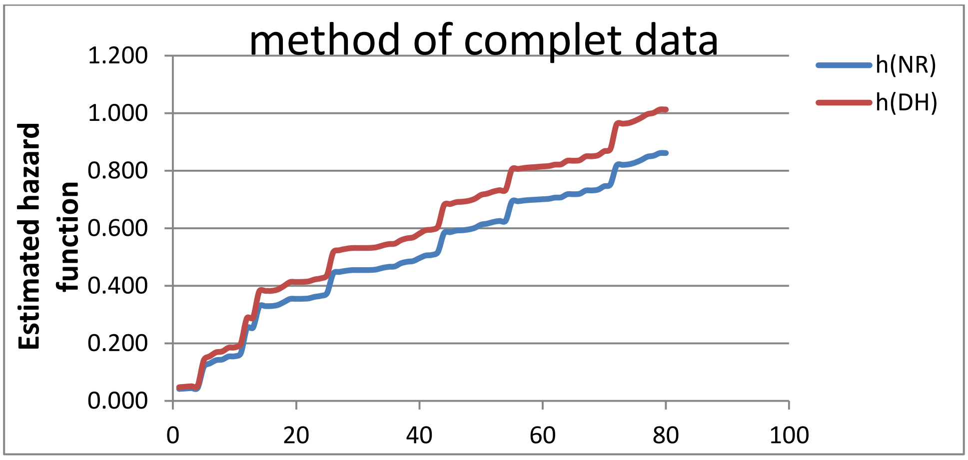


Figure 1: Illustrates the parameter estimates and the hazard function for the estimated parameters using complete data.

Based on Table (4) and using the $RMSEh_{(t,x)}$ comparison metric, it is evident that the Downhill Simplex (DH) algorithm outperforms the Newton-Raphson (NR) algorithm in parameter estimation.

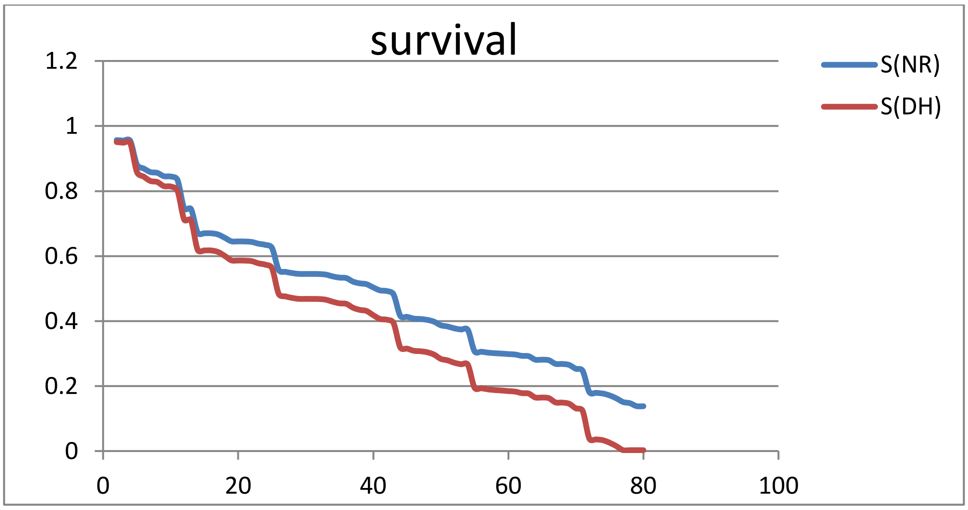
Subsequently, the survival function ($\hat{S}(t)$), was estimated for all cancer patients and the specified time in days. Table (5) and Figure (5) illustrate the estimated survival functions for both methods. From the estimates and the figure, we can observe that the survival function exhibits a decreasing behavior as the hazard function increases over the study period, which aligns with the concept of reliability.

Table 5 : Illustrates the estimation of the survival function $\widehat{S}(t)$

T	S(NR)	S(DH)	T	S(NR)	S(DH)
1	0.958275	0.952344	41	0.494577	0.406607
2	0.956826	0.950662	42	0.492497	0.404296
3	0.955261	0.948863	43	0.482161	0.394157
4	0.954656	0.948088	44	0.416118	0.318842
5	0.881152	0.858335	45	0.413264	0.315522
6	0.869381	0.844382	46	0.4078	0.308846
7	0.85842	0.831029	47	0.406557	0.307367
8	0.856037	0.827965	48	0.403838	0.304192
9	0.845668	0.815164	49	0.398249	0.297076
10	0.844685	0.81397	50	0.387232	0.283957
11	0.83448	0.801371	51	0.383583	0.279304
12	0.746514	0.71217	52	0.377803	0.272027
13	0.744964	0.710318	53	0.374264	0.267366
14	0.671544	0.618789	54	0.373245	0.266176
15	0.670326	0.61748	55	0.307014	0.194992

P-ISSN 2518-5764
E-ISSN 2227-703X

16	0.670174	0.617309	56	0.305879	0.193599
17	0.666524	0.612699	57	0.30306	0.190087
18	0.65648	0.600563	58	0.301343	0.187989
19	0.645463	0.586844	59	0.300148	0.18659
20	0.645161	0.586442	60	0.298609	0.184754
21	0.644904	0.586078	61	0.297389	0.183302
22	0.64353	0.584447	62	0.293186	0.178507
23	0.638135	0.577716	63	0.292225	0.177303
24	0.634571	0.573721	64	0.281412	0.165079
25	0.625167	0.562805	65	0.281354	0.165006
26	0.557165	0.48358	66	0.279731	0.163036
27	0.551569	0.476055	67	0.268537	0.149903
28	0.547867	0.471401	68	0.268188	0.149458
29	0.54523	0.468445	69	0.265331	0.145879
30	0.545166	0.468372	70	0.253412	0.131696
31	0.545049	0.468236	71	0.247386	0.124094
32	0.544859	0.468007	72	0.181385	0.038317
33	0.542993	0.465771	73	0.179402	0.036249
34	0.537485	0.459595	74	0.177314	0.033719
35	0.533811	0.454708	75	0.171517	0.026037
36	0.532413	0.45317	76	0.162387	0.015384
37	0.521434	0.441222	77	0.151135	0.003136
38	0.516204	0.434631	78	0.147581	0.003126
39	0.513563	0.431257	79	0.138505	0.003136
40	0.503467	0.418176	80	0.138274	0.003036



The figure 2: Shows the survival function for the patients.

The Cox model's hazard function for the estimators is shown in the following table and chart.

Table 6 : Shows the hazard function for the model.

t	h(NR)	h(DH)	t	h(NR)	h(DH)
1	0.042	0.048	41	0.505	0.593
2	0.043	0.049	42	0.508	0.596
3	0.045	0.051	43	0.518	0.606
4	0.045	0.052	44	0.584	0.681
5	0.119	0.142	45	0.587	0.684
6	0.131	0.156	46	0.592	0.691
7	0.142	0.169	47	0.593	0.693
8	0.144	0.172	48	0.596	0.696
9	0.154	0.185	49	0.602	0.703
10	0.155	0.186	50	0.613	0.716
11	0.166	0.199	51	0.616	0.721
12	0.253	0.288	52	0.622	0.728
13	0.255	0.290	53	0.626	0.733
14	0.328	0.381	54	0.627	0.734
15	0.330	0.383	55	0.693	0.805
16	0.330	0.383	56	0.694	0.806
17	0.333	0.387	57	0.697	0.810
18	0.344	0.399	58	0.699	0.812
19	0.355	0.413	59	0.700	0.813
20	0.355	0.414	60	0.701	0.815
21	0.355	0.414	61	0.703	0.817
22	0.356	0.416	62	0.707	0.821
23	0.362	0.422	63	0.708	0.823
24	0.365	0.426	64	0.719	0.835
25	0.375	0.437	65	0.719	0.835
26	0.443	0.516	66	0.720	0.837
27	0.448	0.524	67	0.731	0.850
28	0.452	0.529	68	0.732	0.851
29	0.455	0.532	69	0.735	0.854
30	0.455	0.532	70	0.747	0.868
31	0.455	0.532	71	0.753	0.876
32	0.455	0.532	72	0.819	0.962
33	0.457	0.534	73	0.821	0.964
34	0.463	0.540	74	0.823	0.966
35	0.466	0.545	75	0.828	0.974
36	0.468	0.547	76	0.838	0.985
37	0.479	0.559	77	0.849	0.997
38	0.484	0.565	78	0.852	1.002
39	0.486	0.569	79	0.861	1.013
40	0.497	0.582	80	0.862	1.013

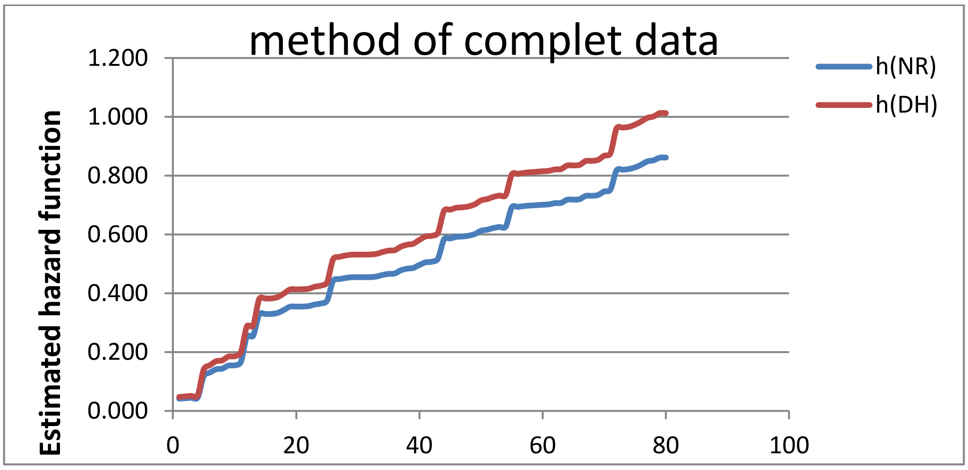


Figure 3: Illustrates the hazard function for the model.

Following that, competitive risks were estimated through the Cumulative Incidence Function (CIF), which is defined as the probability of an event occurring at any time point between the baseline and the specific time. CIF at a time 't' is calculated by the ratio of the individuals who experienced the event (death) divided by the total number of individuals at risk during the time 'ti'. As time progresses, CIF increases from zero to the cumulative proportion of events, as illustrated in the following table.

Table 7: Illustrates the competitive risks for the complete data.

t	$\widehat{H}(t)$	t	$\widehat{H}(t)$	t	$\widehat{H}(t)$	t	$\widehat{H}(t)$
1	0.05	21	0.379999	41	0.558413	61	0.794286
2	0.051764	22	0.381542	42	0.560913	62	0.798401
3	0.053527	23	0.387715	43	0.573259	63	0.799523
4	0.054145	24	0.391831	44	0.653259	64	0.812023
5	0.134145	25	0.404176	45	0.656345	65	0.812089
6	0.14649	26	0.484176	46	0.662518	66	0.813632
7	0.158836	27	0.490426	47	0.66389	67	0.825978
8	0.161336	28	0.494541	48	0.666976	68	0.82639
9	0.173682	29	0.497628	49	0.673149	69	0.829476
10	0.174723	30	0.497703	50	0.685494	70	0.841822
11	0.187069	31	0.497839	51	0.68961	71	0.847994
12	0.267069	32	0.498035	52	0.695783	72	0.927994
13	0.268833	33	0.500092	53	0.701955	73	0.930052
14	0.348833	34	0.506265	54	0.70319	74	0.93211
15	0.350204	35	0.51038	55	0.78319	75	0.938283
16	0.350367	36	0.511924	56	0.784579	76	0.950628
17	0.354482	37	0.524269	57	0.787704	77	0.962974
18	0.366828	38	0.530442	58	0.789468	78	0.967089
19	0.379328	39	0.533567	59	0.790839	79	0.979435
20	0.379666	40	0.545913	60	0.792897	80	0.979685

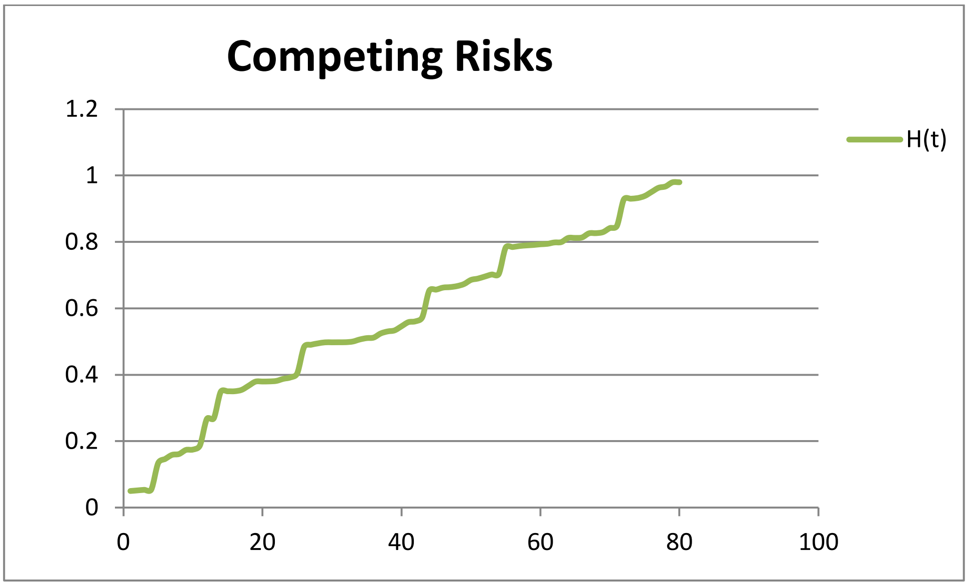


Figure 4: Illustrates the cumulative hazard function for competitive risks.

Through Table (7) on competitive risks and compared with Table (6) on the risk function for the Cox model and its dependent figures, we notice that competitive risks are greater than the risk function for the estimator. Newton Raphson (NR) indicates the effect of common variables in increasing risks function on the hazard function, while competitive risks converge with the hazard function of the estimator and Downhill (Dh).

4.Conclusion:

1.The Down-Hill algorithm demonstrated its superiority in estimating the Cox model for both cases of parameters that were assumed and for various models of means, standard deviations, and different sample sizes in the simulation.

2.The preference for the methods of the applied side coincided with the simulation. The results of the applied side showed that the Downhill algorithm is also the best in estimation.

3.Simulation experiments show that the statistical standard error (RMSE) of the Cox model decreases as the sample size increases.

4.Through the results of Competing risks for real data, I explained the effect of common variables in increasing the risk over the hazard function of the Cox model in some estimators and their convergence in other estimators.

5.From the applied aspect, we notice that the longer the time, the greater the risk for the model, which leads to a decrease in the survival function.

Authors Declaration:

Conflicts of Interest: None

-We Hereby Confirm That All The Figures and Tables In The Manuscript Are Mine and Ours. Besides, The Figures and Images, Which are Not Mine, Have Been Permitted Republication and Attached to The Manuscript.

- Ethical Clearance: The Research Was Approved By The Local Ethical Committee in The University.

References:

1.Andrade, C. 2023. "Survival analysis, Kaplan-Meier curves, and Cox regression: basic concepts". *Indian Journal of Psychological Medicine*, 02537176231176986. 45(4), pp.434- 435.

2.Teoh, S. H. 2008. "Using Kaplan Meier and cox regression in survival analysis: An example".*Teoh Sian Hoon. ESTEEM Academic Journal*, 4(2), pp.3–14 . .

3.Smith, T., and Smith, B. 2003. "Kaplan Meier and Cox proportional hazards modeling: hands on survival analysis". *SAS® Users Group International Proc. Seattle, Washington*. 6(2), pp.1–14.
4.Tang, Z., Zhou, C., Jiang, W., Zhou, W., Jing, X., Yu, J., and Sheng, B. 2013. "Analysis of significant factors on cable failure using the cox proportional hazard model". *IEEE Transactions on Power Delivery*, 29 (2), pp.951–957.
5.Scheike, T. H., and Sun, Y. 2007." Maximum likelihood estimation for tied survival data under Cox regression model via EM-algorithm". *Lifetime data analysis*, 13(4), pp.399-420 .
6.Thackham, M., and Ma, J. 2020. "On maximum likelihood estimation of the semi-parametric Cox model with time-varying covariates". *Journal of Applied Statistics*, 47(9), pp.1511–1528 .
7.Yalçınkaya, A., Şenoğlu, B., and Yolcu, U. 2018." Maximum likelihood estimation for the parameters of skew normal distribution using genetic algorithm". *Swarm and Evolutionary Computation*, 38 (16), pp.127–138.
8.Akram, S., and Ann, Q. U. 2015." Newton Raphson method". *International Journal of Scientific & Engineering Research*, 6(7), pp.1748–1752.
9.Casella, F., and Bachmann, B. 2021. "On the choice of initial guesses for the Newton-Raphson algorithm". *Applied Mathematics and Computation*, 3(18), pp.1–38.
10. Fajfar, I., Puhan, J., and Bűrmen, Á. 2017. "Evolving a Nelder–Mead algorithm for optimization with genetic programming". *Evolutionary computation*, 25(3), pp.351–373.
11. Galántai, A. 2021. "A convergence analysis of the Nelder-Mead simplex method". *Acta Polytech. Hung*, 18(5), pp.93–105.
12. Mehta, V. K. 2020. "Improved nelder–mead algorithm in high dimensions with adaptive parameters based on chebyshev spacing points". *Engineering Optimization*, 52(10), pp.1814 – 1828.
13. Fan, J., Lin, H., and Zhou, Y. 2006." Local partial-likelihood estimation for lifetime data". 34(1), pp.290–325.
14. Sinha, D., Ibrahim, J. G., and Chen, M. H. 2003. "A Bayesian justification of Cox's partial likelihood". Biometrika, 90(3), pp.629-641.
15. Cox, D. R. 1975. "Partial likelihood. Biometrika", 62(2), pp.269-276.
16. Groeneboom, P., Maathuis, M. H., and Wellner, J. A. 2008. "Current status data with competing risks: Limiting distribution of the MLE". *Annals of statistics*,.36 (3), pp.1064-1089.
17. Kalbfleisch, J. D., and Prentice, R. L. 1973. "Marginal likelihoods based on Cox's regression and life model". *Biometrika*, 60(2), pp.267-278.
18. Sildnes, B. 2015. "Modelling of Dependent Competing Risks and Semi-Competing Risks by Means of First Passage Times of Gamma Processes". (*Master's thesis*, *NTNU*), 85(6), pp.1-258.
19. Hudgens, M. G., Li, C., and Fine, J. P. 2014. "Parametric likelihood inference for interval censored competing risks data". *Biometrics*, 70(1), pp.1-9.
20. Mao, L., and Lin, D. Y. 2017. "Efficient estimation of semiparametric transformation models for the cumulative incidence of competing risks". *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 79(2) , pp.573-587.
21. Thackham, M., and Ma, J. 2020. "On maximum likelihood estimation of the semi-parametric Cox model with time-varying covariates". *Journal of Applied Statistics*, 47(9) , pp.1511-1528.
22. Zeng, D., Mao, L., and Lin, D. Y. 2016. "Maximum likelihood estimation for semiparametric transformation models with interval-censored data". *Biometrika*, 103(2),pp. 253-271.
23. Guerrero, A., Del Campo, L., Piscaglia, F., Scheiner, B., Han, G., Violi, F., and Girleanu, I. 2023. "Anticoagulation improves survival in patients with cirrhosis and portal vein thrombosis": *The IMPORTAL competing-risk meta-analysis. Journal of Hepatology* , 79 (1), pp.69-78.

Journal of Economics and Administrative Sciences
2024; 30(140), pp. 450-465
P-ISSN 2518-5764
E-ISSN 2227-703X

استعمال طريقة الامكان الاعظم في تقدير نموذج كوكس لبيانات المخاطرة التنافسية مع التطبيق

ابراهيم خليل حسين
جامعة بغداد / كلية الادارة والاقتصاد / قسم الاحصاء
العراق
ibrahim.khaleel2101m@coadec.uobaghdad.edu.iq

انتصار عريبي فدعم
جامعة بغداد /كلية الادارة والاقتصاد / قسم الاحصاء
العراق
entsar.arebe@coadec.uobaghdad.edu.iq

Received:26/9/2023 Accepted:20/12/2023 Published Online First: 30 /4/ 2024

مستخلص البحث:
تناولت العديد من البحوث و الدراسات السريرية حدوث الفشل (الموت) نتيجة عوامل اخرى (خارجية) اي مخاطر اضافية قد تحدث تنافس الحدث قيد الدراسة , تتم احالة البيانات الناتجة الى البيانات الكاملة للمخاطر التنافسية والتي تتأثر باختلاف الوقت , وان وقت حدوث هذه المخاطر معلوم بالضبط , اي يتم ملاحظة اوقات الفشل لكل المشاهدات بصوره مؤكده , والتي تم تقدير تأثيرها على دالة الخطر اعتمادا على نموذج كوكس للمخاطر النسبية والذي تم تقديره من خلال الجزء المعلمي المتمثل بطريقة الامكان الاعظم الجزئي واستعمال الخوارزميات العددية لتقدير المعلمات والمتمثلة بتأثير المتغيرات على دالة الخطر لكوكس والجزء اللامعلمي المتمثل بتقدير تأثير الوقت على دالة الخطر من خلال صيغة كابلان ماير وحساب المخاطر التنافسية عن طريق دالة الخطر التراكمي . حيث تم تطبيق هذه الاساليب على البيانات التجريبية من خلال عمليات محاكاة واسعة النطاق لأحجام ومعالم مختلفو ولعدة نماذج من الاوساط الحسابية والانحرافات المعيارية وكذالك التطبيق على بيانات حقيقية من خلال عينة بحجم (80) من الاشخاص المصابون بسرطان الثدي , من خلال تحليل كل من نتائج المحاكاة والبيانات الحقيقية بينت النتائج التي توصلنا اليها ان خوارزمية داون هيل تتفوق على خوارزمية نيوتن رافسون من حيث دقة التقدير والكفاءة اعتمادا على المعيار الاحصائي للمقارنة جذر متوسط مربع الخطأ , اضافة لذلك اوضحت المخاطر التنافسية تأثير المتغيرات المشتركة في زيادة المخاطر التنافسية عن دالة الخطر لنموذج كوكس للمقدر نيوتن رافسون في حين اوضحت تقاربها من دالة الخطر لنموذج كوكس بالنسبة للمقدر داون هيل .

نوع البحث: ورقة بحثية.
الكلمات الرئيسة: دالة الخطر ; البيانات الكاملة ; نموذج كوكس للمخاطر النسبية ; الامكان الاعظم الجزئي ; خوارزمية نيوتن رافسون ; خوارزمية داون هيل .