Australian Journal of Basic and Applied Sciences, 7(10): 602-607, 2013
ISSN 1991-8178

Analysis of Survival Data Using Cox Model (Continuous Type)

Khawla Mustafa Sadiq

Department of Mathematics, Education College, Mosul University, Iraq

Abstract: We outline the type of data and the process in data collection that defines such experimental situation using a fish experiment, where studying the effect of treatment combinations on the survival times of fish in aquarium water was desired. In an experiment presented by pierce, steware and kopoecky (1978), fish were subjected to three levels of zinc concentration in aquarium water, and approximate times to death were observed it was desired to study the effect of either one or two week's acclimation in the test aquaria before introduction of the zinc. There were initially two tanks for each of the treatment combinations. The experiment was a 2×3 factorial for the treatment combinations. The $2\forall 3$ treatment combinations were assigned to tanks in a completely randomized design. From this point on wards we use CRD to designate this design. The experiment was carried for 7 days and mortality was observed on daily basis. Three hundred fish were randomized to 12 tanks, 25 fish to each tank. The 2×3 treatment combinations were assigned so that 2 tanks received each treatment.

Key words: survival analysis, Cox model, multinomial distribution, marginal likelihood, split plot.

INTRODUCTION

Several experimental situation give rise to analyzing time to response on observation units (survival data) using split plot in time models. The general structure of such experiments is that the observation of time of occurrence of an event called a death, failure, or response is of interest.

The observational units are grouped into whole units and the treatments are randomized to whole units. If time to the occurrence of an event T is a continuous random variable then whole units would be considered as sub samples, if time response was grouped into intervals in the above setting, then the sufficient statistics in this case would be the counts of observed occurrences of an event (number of death, failure) within intervals . The experiment can then be viewed as a split plot over time where time intervals (periods) are subunits and whole units would be the same as in continuous time setting, and the response variable is some function of the counts. For the split plot over time model we are interested in estimating survival curves rather than means for the usual structure of split plot model.

Cox (1972) considered the analysis of censored failure time. He suggested regression model for the failure time T of an individual when values of the one or more explanatory variable were available. For T continuous , the hazard is given by: Λ (t , z) = λ_0 (t) exp (β', Z) which is known as the proportional hazard function with β being the vector of the unknown parameters, and λ_0 (t) is the underlying hazard function when z=0 for T discrete, the logistic model was suggested. A conditional likelihood and maximum likelihood estimates was obtain, However, Cox proportional hazard regression model does not handle grouped survival data or large data with many ties (many individuals failed at the same time) Kalbfleisch & Prentice (1973) obtained a marginal likelihood for the regression parameters by restricting the class of model presented by Cox to those that possessed a strictly monotone survivor function or equivalently, to those for which the hazard function λ_0 (t) was not identically zero over an open interval.

Definitions:

Let us denote the points defining the time interval by: $0 = t_0 < t_1 < t_2 < t_3 < t_4 < t_5 < t_6 < t_7$. $f(x) = \lambda_0\, e^{-\lambda t}$ λ_0 is hazard function. The number of failures or deaths in each day would be the number of failures or deaths in time intervals $(t_{k-1}, t_k]$ for K=1,2,…,7 Also , define:

n_{ijk}: number assigned (at risk) to trt i, time interval k and tank j.

s_{ijk} :number of survivors during interval k on trt i and tank j.

r_{ijk} :number of failures on trt i during interval k for tank j.

p_{ijk}: conditional probability that a unit on trt i fails in time interval k given that it survived (k-1) time intervals for a given tank j.

q_{ijk}: conditional probability that a unit on trt i survives time interval k given that it survived k-1 time intervals for a given tank j, where $q_{ijk} = 1 - p_{ijk}$.

Corresponding Author: Khawla Mustafa Sadiq, Department of Mathematics, Education College, Mosul University, Iraq
E-mail: khawlamus@yahoo.com, Mob: 009647710747316

The Selected Model:

Assume that tank effects increase or decrease the survivals, ie. there is tank variability involved, since treatment combinations were applied to main units (tanks). Also failure time T is a discrete random variable since time responses were grouped into intervals. The response for the discrete setting would be some function of the number of deaths or the number of survivors. This will give us a split plot in time where subplot units are time intervals. Failure time variability will arise from the fact that 25 fish were randomly assigned to each tank, as we can see in table (1). Assuming that condition on being in the same tank survival times of different fish are independent, then the model to be considered is:

Response = $\mu + \alpha_i + \varepsilon_{ij} + B_k + (\alpha B)_{ik} + \delta_{ijk} \quad i = 1,2,\dots,I,\ j = 1,2,\dots,J \text{ and } k = 1,2,\dots,k$

μ: is an overall mean.

α_i: is treatment combination i effect,

ε_{ij}: is main unit variability (tank variability) With: $E(\varepsilon_{ij}) = 0$, $E(\varepsilon_{ij}\ \varepsilon'_{ij}) = \sigma_\varepsilon^2 \quad where\ j = j'$
$= 0 \quad where\ j \neq j'$

β_k : is the subplot treatment or time interval effect

$(\alpha\ \beta)_{ik}$: is the interaction between treatment and time interval .

δ_{ijk} : is the variability due to the different fish in each tank with: $E(\delta_{ijk}) = 0$, $E(\epsilon_{ij}\ , \delta_{ijk}) = 0$, $and\ \ E(\delta_{ijk}\,\delta'_{ijk}) = \sigma^2_{\delta_{ijk}} \quad where\ k = k'$
$= 0 \quad where\ k \neq$ k'

The response of the above model will depend on the model assumed for the hazard function for time interval k and trt i. The hazard function $\lambda_i(t_k)$ is the conditional probability of failing in an interval given surviving until that interval. The choice for the response is $f(\hat{q}_{ijk})$. Two possible choices for this function that will be considered are: $f(\hat{q}_{ijk}) = \log(1 - \log \hat{q}_{ijk})$ and $f(\hat{q}_{ijk}) = \log(\hat{q}_{ijk})$. These response are derived from continuous random models. From this point onwards we use log(x) to denote $\log_e(x)$.

The Distribution of the Number Survivor:

Individuals at risk during time interval k may fail, be censored, or survive to the start of the following time period. Assuming that there is no censoring, the observed number at risk for time interval k on a given trt i and tank j is n_{ijk}, and the number of individuals failing is r_{ijk}. Define $n_{ijk\,(k+1)} = n_{ijk} - r_{ijk}$, which is denoted by s_{ijk} (the number of individuals surviving interval k). Thus, individuals surviving interval k will be individuals at risk for the next time interval, i.e., $s_{ijk} = n_{ijk\,(k+1)}$. For a given trt i , tank j, and k time intervals, number of deaths or failures $r_{ij1}, r_{ij2}, \dots\dots\dots r_{ijk}$ in time intervals (t_o , t_1], (t_1 , t_2], …, (t_{k-1} , t_k] with t_o=0 among n_{ij1}. Starters, follow a multinomial distribution with probability function: (7)

$$\text{pr}(r_{ij1}, r_{ij2}, \dots, r_{ijk}/\boldsymbol{\varepsilon}_{ij}) = \frac{n_{ij}!}{r_{ij1}!\, r_{ij2}! \dots r_{ij(k+1)}!} \prod_{k=1}^{k+1} \pi_{ijk}^{r_{ijk}}$$, Where: $r_{ij1} + r_{ij2} + \cdots + r_{ij\,(k+1)} = n_{ij1}$ and $\pi_{ij1} +$

$\pi_{ij2} + \cdots + \pi_{ij\,(k+1)} = 1$.

Now define: $\boldsymbol{p_{ijk}} = \prod_{l=1}^{k} \boldsymbol{q_{ijl}}$ is the probability an individual on trt i and in tank j survives beyond interval k, $\pi_{ijk} = p_{ij\,(k-1)} - p_{ijk}$ is the probability an individual fails in interval k for a given tank j on trt i, q_{ijk} is the conditional probability an individual on trt i and in tank j survives beyond interval k given that it survives beyond interval k-1, where $q_{ijk} = p_{ijk}/p_{ij\,(k-1)}$, $p_{ijk} = 1 - q_{ijk}$ is the conditional probability an individual on trt I and in tank j fails in interval k given that it survives beyond interval k-1, and $r_{ij\,(k+1)} = s_{ijk}$ is the number of individuals surviving at end of study.

Therefore, we have (6) $\pi_{ijk} = p_{ij\,(k-1)} - p_{ijk} = q_{ij1} q_{ij2} \dots . q_{ij\,(k-1)} p_{ijk}$ For k=1,2,…,k. the likelihood function for the multinomial distribution is given by:

$$\text{pr}(r_{ij1}, r_{ij2}, \cdots, r_{ijk}/\boldsymbol{\varepsilon}_{ij}) \boldsymbol{\alpha} \prod_{k=1}^{k+1} \pi_{ijk}^{r_{ijk}} \boldsymbol{\alpha} \left\{ \prod_{k=1}^{k} \left(q_{ij1} \cdots . q_{ij(k-1)} p_{ijk}\right)^{r_{ijk}} \right\} . \left\{ \left(q_{ij1} \cdots . q_{ij(k-1)} . q_{ijk}\right)^{r_{ij(k+1)}} \right\}$$

$$\alpha\ \ p_{ij1}^{r_{ij1}} p_{ij2}^{r_{ij2}} p_{ij3}^{r_{ij3}} \dots p_{ijk}^{r_{ijk}} \ \ . \ \ q_{ij1}^{r_{ij2} + r_{ij3} + \cdots + r_{ij\,(k+1)}} \ \ . \ \ q_{ij2}^{r_{ij3} + r_{ij4} + \cdots + r_{ij\,(k+1)}} \dots q_{ijk}^{r_{ij\,(k+1)}}$$

Recall that: $r_{ij1} + r_{ij2} + \cdots + r_{ij\,(k+1)} = n_{ij1}$, and $n_{ijk} = n_{ij1} - r_{ij1} - r_{ij2} - \cdots - r_{ij\,(k-1)}$ for k=1,2,…,k+1. Therefore the likelihood is proportional to:

$$pr(r_{ij1}, r_{ij2}, \dots, r_{ijk}/\varepsilon_{ij}) \ \alpha\ \ p_{ij1}^{r_{ij1}} q_{ij1}^{n_{ij1} - r_{ij1}} \ \ p_{ij2}^{r_{ij2}} q_{ij2}^{n_{ij1} - r_{ij1} - r_{ij2}} \ \ \dots \ \ p_{ijk}^{r_{ijk}} q_{ijk}^{n_{ij1} - r_{ij1} - r_{ij2} - \cdots - r_{ijk}}$$

$$\alpha\ \ p_{ij1}^{r_{ij1}} q_{ij1}^{n_{ij1} - r_{ij1}} \ \ p_{ij2}^{r_{ij2}} q_{ij2}^{n_{ij2} - r_{ij2}} \ \ \dots \ \ p_{ijk}^{r_{ijk}} q_{ijk}^{n_{ijk} - r_{ijk}}$$

$$\alpha\prod_{k=1}^{k} p_{ijk}^{r_{ijk}} q_{ijk}^{n_{ijk}-r_{ijk}}$$

$$\alpha\prod_{k=1}^{k} q_{ijk}^{s_{ijk}} \left(1-q_{ijk}\right)^{n_{ijk}-s_{ijk}}$$

Therefore, conditioning on n_{ijk}, the number of survivors s_{ijk} in time interval k on trt i and a given tank j is distributed as a binomial random variable with parameters n_{ijk} and q_{ijk}. Furthermore, the covariance between s_{ijk} and $s_{ij\,k'}$, is zero. Also, the mean and variance of $\hat{q}_{ijk}$ given that n_{ijk} is fixed by its observed number and for a given tank are given by

$E\left(\hat{q}_{ijk}/n_{ijk},\varepsilon_{ij}\right) = q_{ijk}$, and $\mathrm{Var}\left(\hat{q}_{ijk}/n_{ijk},\varepsilon_{ij}\right) = p_{ijk}\,q_{ijk}/n_{ijk}$, respectively.

Now for $k < k'$, assuming that $n_{ijk} > 0$ we have

$$E\left(\left(\hat{q}_{ijk}-q_{ijk}\right)/\hat{q}_{ijk}',\varepsilon_{ij},n_{ijk}\right) = E_{s_{ijk}}\left(\left(\frac{s_{ijk}}{n_{ijk}}-q_{ijk}\right)/\hat{q}_{ijk}',\varepsilon_{ij},n_{ijk}\right)$$

$$= E_{n_{ijk}}\left\{E_{s_{ijk}}/n_{ijk}\left(\left(\frac{s_{ijk}}{n_{ijk}}-q_{ijk}\right)/\hat{q}_{ijk}',\varepsilon_{ij},n_{ijk}\right)\right\} = E_{n_{ijk}}\left\{\left(0/\hat{q}_{ijk}',\varepsilon_{ij},n_{ijk}\right)\right\} = 0$$

Hence, for k<k'

$\mathrm{cov}\left(\hat{q}_{ijk},\hat{q}_{ijk}'/\varepsilon_{ij},n_{ijk}\right) = E_{\hat{q}_{ijk}'}\left(\left(\hat{q}_{ijk}'-\hat{q}_{ijk}\right)E\left(\left(\hat{q}_{ijk}-q_{ijk}\right)/\hat{q}_{ijk}',\varepsilon_{ij},n_{ijk}\right)\right) = 0$

for a given tank j and a fixed risk set (n_{ijk}) we have the equation before $s_{ijk}/\varepsilon_{ij}, n_{ijk} \sim$ binomial (n_{ijk}, q_{ijk}), and

$$\mathrm{cov}\left(s_{ijk},s_{ij\,k'}/\varepsilon_{ij},n_{ijk},n_{ij\,k'}\right) = 0$$

Thus, for a large sample size the asymptotic distribution is given by

$s_{ijk}/\varepsilon_{ij},n_{ijk} \sim normal\left(n_{ijk}\,q_{ijk}, n_{ijk}\,q_{ijk}\left(1-q_{ijk}\right)\right)$, and

$\hat{q}_{ijk} = \frac{s_{ijk}}{n_{ijk}}/\varepsilon_{ij},n_{ijk} \sim normal\left(q_{ijk}, q_{ijk}\left(1-q_{ijk}\right)/n_{ijk}\right)$

Therefore, equal variance structure of q_{ijk}'s would be inappropriate since these variances depend on $\hat{q}_{ijk}$'s which may vary over time, and the fact that the risk sets decrease over time.

Estimation of Survival Function:

Our full model is given by $\underset{\sim}{y} = \boldsymbol{x}\,\underset{\sim}{\boldsymbol{\beta}} + \underset{\sim}{\boldsymbol{u}}$, where cov (y) = v and v is a block diagonal variance – covariance matrix involving both $\sigma^2_{\delta_{ijk}} = \frac{1-\hat{q}_{ijk}}{n_{ijk}\;q_{ijk}\;(\log q_{ijk})}$ and $\hat{\sigma}^2_{\epsilon}$ = [R($\underset{\sim}{\boldsymbol{b}},\underset{\sim}{\boldsymbol{c}}$) - R(b) - tr[c_1] + tr[c_2]] / tr [c] . $\underset{\sim}{\boldsymbol{\beta}}$ is the vector of unknown parameters to be estimated, x is a design matrix of known constants, and $\underset{\sim}{\boldsymbol{y}}$ is a vector of transformed values of the observed $\hat{q}_{ijk}'s$. The function of these $\hat{q}_{ijk}'s$ that was considered in the analysis is given by: $y_{ijk} = \log(-\log\hat{q}_{ijk})$. If the estimate of $\hat{\sigma}^2_{\epsilon}$ is small, it will be considered throughout after constructing $\hat{v}$, a weighted least squares procedure is used to fit the full model, where $\hat{B} = (x'\hat{v}^{-1}x)^{-1}x'\hat{v}^{-1}$ *and* $\widehat{cov}\left(\hat{\beta}\right) = (x'\hat{v}^{-1}x)^{-1}$. Since the best fitting model has no time by treatments interactions then the proportional hazed for continuous time setting is appropriate. We use for the analysis of deviation X^2 and p-values.

Table 1: Observed number of deaths.

Acclimation Time	One week						Two weeks					
Zinc concentration	Lo		Med		Hi		Lo		Med		Hi	
Tank / Day Mortality	1	2	1	2	1	2	1	2	1	2	1	2
1	0	0	0	0	0	0	0	0	0	0	0	0
2	1	2	3	0	1	1	0	0	1	0	3	0
3	5	7	7	10	12	10	9	4	12	9	12	12
4	7	4	9	7	7	8	4	4	5	3	3	7
5	1	2	0	5	4	3	0	0	3	2	2	2
6	0	0	0	1	0	1	0	0	0	0	1	0
7	0	0	0	1	1	1	0	0	0	0	0	0

Table 2: Risk set table.

Acclimation Time	One week						Two weeks					
Zinc concentration	Lo		Med		Hi		Lo		Med		Hi	
Tank / Interval/n_{ijk}	1	2	1	2	1	2	1	2	1	2	1	2
1	25	25	25	25	25	25	25	25	25	25	25	25
2	24	23	22	25	24	24	25	25	24	25	22	25
3	19	16	15	15	12	14	16	21	12	16	10	13
4	12	12	6	8	5	6	12	17	7	13	7	6
5	11	10	6	3	1	3	12	17	4	11	5	4
6	11	10	6	2	1	2	12	17	4	11	4	4
7	11	10	6	2	0	1	12	17	4	11	4	4

Table 3: Values of $\hat{q}_{ijk}$.

Acclimation Time	One week						Two weeks					
Zinc concentration	Lo		Med		Hi		Lo		Med		Hi	
Tank / Interval/$\hat{q}_{ijk}$	1	2	1	2	1	2	1	2	1	2	1	2
1	0.960	0.920	0.880	0.980	0.960	0.960	0.980	0.980	0.960	0.980	0.880	0.980
2	0.792	0.696	0.682	0.600	0.500	0.583	0.690	0.840	0.500	0.640	0.455	0.520
3	0.632	0.950	0.400	0.533	0.417	0.420	0.750	0.810	0.583	0.813	0.700	0.962
4	0.917	0.833	0.917	0.375	0.200	0.500	0.958	0.971	0.571	0.816	0.714	0.667
5	0.955	0.950	0.917	0.667	0.500	0.667	0.958	0.971	0.875	0.955	0.800	0.875
6	0.955	0.850	0.917	0.500	0.500	0.500	0.958	0.971	0.875	0.955	0.875	0.875
7	0,955	0.950	0.917	0.500	0.500	0.500	0.958	0.971	0.875	0.955	0.875	0.875

Table 4: Estimates of Binomial Variances And Values of The Response Variable.

Accl	Conc	tank	Time	r_{ijk}	n_{ijk}	s_{ijk}	$\hat{\sigma}^2_{\delta_{ijk}}$	y_{ijk}
1	1	1	1	1	25	24	0.7415	3.1985
1	1	1	2	5	24	19	0.1728	-1.4559
1	1	1	3	7	19	12	0.19855	-0.779
1	1	1	4	1	12	11	0.7531	-2.4459
1	1	1	5	0	11	11	2.08315	-3.0782
1	1	1	6	0	11	11	2.08315	-3.0782
1	1	1	7	0	11	11	2.08315	-3.0782
1	1	2	1	2	25	23	0.74745	-2.4843
1	1	2	2	7	23	18	0.1728	-1.015
1	1	2	3	4	16	12	0.19855	-1.2459
1	1	2	4	2	12	10	0.7531	-1.6998
1	1	2	5	0	10	10	2.08315	-2.9702
1	1	2	6	0	10	10	2.08315	-2.9702
1	1	2	7	0	10	10	2.08315	-2.9702
1	2	1	1	3	25	22	1.1878	-2.057
1	2	1	2	7	22	15	0.12345	-0.9604
1	2	1	3	9	15	6	0.1333	-0.0874
1	2	1	4	0	6	6	1.1138	-2.4459
1	2	1	5	0	6	6	1.5129	-2.4459
1	2	1	6	0	6	6	2.012	-2.4459
1	2	1	7	0	6	6	1.5258	-2.4459
1	2	2	1	0	25	25	1.1878	-3.9019
1	2	2	2	10	25	15	0.12345	-0.6717
1	2	2	3	7	15	8	0.1333	-0.4633
1	2	2	4	5	8	3	1.1138	0.0194
1	2	2	5	1	3	2	1.5129	-0.904
1	2	2	6	0	2	2	2.012	-1.2459
1	2	2	7	1	2	1	1.5258	-0.3665
1	3	1	1	1	25	24	0.9804	-3.1985
1	3	1	2	12	24	12	0.0946	-0.3665
1	3	1	3	7	12	5	0.1425	0.1339
1	3	1	4	4	5	1	0.3279	0.4759
1	3	1	5	0	1	1	1.5479	-0.3665
1	3	1	6	1	1	0	1.5609	-0.3665
1	3	1	7	0	0	0	2.0812	-0.3665
1	3	2	1	1	25	24	0.9804	-3.1985
1	3	2	2	10	24	14	0.0946	-0.611
1	3	2	3	8	14	6	0.1425	-0.1669
1	3	2	4	3	6	3	0.3274	-0.3665
1	3	2	5	1	3	2	1.5479	-0.904
1	3	2	6	1	2	1	1.5609	-0.3665
1	3	2	7	0	1	1	2.0812	-0.3665
2	1	1	1	0	25	25	2.0408	-3.9019

2	1	1	2	9	25	16	0.1818	-0.8068
2	1	1	3	4	16	12	0.2516	-1.2159
2	1	1	4	0	12	12	1.9909	-3.1487
2	1	1	5	0	12	12	1.9909	-3.1487
2	1	1	6	0	12	12	1.9909	-3.1487
2	1	1	7	0	12	12	1.9909	-3.1487
2	1	2	1	0	25	25	2.0408	-3.9019
2	1	2	2	4	25	21	0.1818	-1.7467
2	1	2	3	4	21	17	0.2516	-1.5572
2	1	2	4	0	17	17	1.9909	-3.5258
2	1	2	5	0	17	17	1.9909	-3.5258
2	1	2	6	0	17	17	1.9909	-3.5258
2	1	2	7	0	17	17	1.9909	-3.5258
2	2	1	1	1	25	24	1.5106	-3.1985
2	2	1	2	12	24	12	0.0999	-0.3665
2	2	1	3	5	12	7	0.27	-0.617
2	2	1	4	3	7	4	0.421	-0.5792
2	2	1	5	0	4	4	2.0231	-2.0134
2	2	1	6	0	4	4	2.0231	-2.0134
2	2	1	7	0	4	4	2.0231	-2.0134
2	2	2	1	0	25	25	1.5106	-3.9019
2	2	2	2	9	25	16	0.0999	-0.8068
2	2	2	3	3	16	13	0.27	-1.5749
2	2	2	4	2	13	11	0.421	-1.7883
2	2	2	5	0	11	11	2.0231	-3.0782
2	2	2	6	0	11	11	2.0231	-3.0782
2	2	2	7	0	11	11	2.0231	-3.0782
2	3	1	1	3	25	22	1.1871	-2.057
2	3	1	2	12	22	10	0.0871	-0.2389
2	3	1	3	3	10	7	0.2436	-1.0309
2	3	1	4	2	7	5	0.5058	-1.0881
2	3	1	5	1	5	4	1.5052	-1.4999
2	3	1	6	0	4	4	2.0064	-2.043
2	3	1	7	0	4	4	2.0064	-2.0134
2	3	2	1	0	25	25	1.1877	-3.9019
2	3	2	2	12	25	13	0.0871	-0.4248
2	3	2	3	7	13	6	0.2436	-0.2585
2	3	2	4	2	6	4	0.5057	-0.904
2	3	2	5	0	4	4	1.5052	-2.0134
2	3	2	6	0	4	4	2.0064	-2.0134
2	3	2	7	0	4	4	2.0064	-2.0134

Table 5: Analysis of deviation for the full model.

Source	d.f	X^2	p-value
Accl	1	11.8354	< 0.005
Conc	2	19.7773	< 0.005
Accl × Conc	2	3.5608	0.290
Time	6	18.3412	0.005
Acc × time	6	7.6430	0.380
Con × Time	12	8.1059	0.790
Codness of fit	48	20.9127	0.995

Table 6: Estimate of $\underset{\sim}{\beta}$.

Parameter	Estimate	S.e
M	-1.741860	0.11862
a_1	-0.561034	0.11291
c_1	0.103323	0.10440
c_2	0.133520	0.07461
T_1	-1.432720	0.28642
T_2	0.945690	0.14188
T_3	1.016890	0.15675
T_4	0.574620	0.22061
T_5	-0.362336	0.34149
T_6	-0.38847	0.3579

RESULTS AND DISCUSSION

We used a real data set for the model using the SAS system and the tables from 1 to 7 show the applications. Since the 50 fish were randomly assigned to each treatment combination with 2 tanks for each treatment the 25 fish were assigned to each tank. Thus the number at risk for the first time interval is 25 fish and the size of this risk set, n_{ijk}, decrease as time advances. For the no-censoring case, the number at risk for time interval k would be the number at risk for time interval k-1 minus the number of deaths for time interval k-1. Therefore, table 2 represents risk sets (values of n_{ijk}). Now let us define $\hat{q}_{ijk} = s_{ijk}/n_{ijk}$, where $s_{ijk} = n_{ijk} - r_{ijk}$ assuming no censoring. Table 3 represents the values of $\hat{q}_{ijk}$. If $s_{ijk} = n_{ijk}$ the use $s_{ijk} - 0.5$, if $s_{ijk} = 0$ then use 0.5 . It should be mentioned here that in order to avoid having values for $\hat{q}_{ijk}$ as one or zero adjusted survival, s_{ijk} (AD) can be used instead of the s_{ijk} as follow Grizzle, starmer and Koch (1969), and suggest that s_{ijk} can be replaced by 1/k, where k is the number of time intervals. The value of $\hat{\sigma}_{\epsilon}^2$ is 0.0095.

Conclusions:

1- For the above analysis we conclude to εij=0, the effect of the acclimation time was important in explaining the data. For the first two time intervals there was practical no different in survival rates between. Acclimation time of one week and two weeks, fish under two weeks acclimations survived better than those with one week acclimation time in the sense that effect become greater with time. This suggests it is better to collect the data (count the number of deaths) after a period of at lead three days. There was also an effect due to zinc concentration which indicates fish survive better with low levels of zinc concentration that for higher level.

2- When we are using Cox and modeling it, we must add a random variable which include all the variates which is occur as a result from transfer the model from the discrete type to continuous type or any variables, name ϵ_{ij}.

3- Using any things which is available instead of any chemical intersection.

REFERENCES

Elandt-Johnson, R.C., 1980. Some Prior Distributions in Survival Analysis: A Critical Insight on Relationship Derived from cross-sectional data. J.R. statist. Soc., 42: 96-106.

Grizzle, J.E., C.F. Starmer and E. Koch, 1969. Analysis of Categorical Dat by Linear Models. Biometrics, 25: 489-584.

Hartley, H.O. and K.S.E. Jayatillake, 1973. Estimation of Linear Models with Unequal Variances. J. Amer statistics. Assoc., 68: 189-192.

Henderson, C.R., 1953. Estimation of Varian and Covariance Components. Biometrics, 9: 226-252.

Kaplan, E.L. and P. Meier, 1958. Non Parametrical Estimation from Incomplete Observations. J. Amer. Statist. Assoc., 53: 457-481.

Kay, R., 1977. Proportional Hazard Regression Models and the Analysis of Censored Survival Data. Appl. Statis., 26: 227-237.

Krane, S.A., 1963. Analysis of Survival Data by Regression Techniques. Techno metrics, 5: 161-174.

Wassertheil-Smoller, Sylvia, 2010. "Biostatistics and Epidemiology: A primer For Health Biomedical Professional" Third edition.

World Scientific Review, 2006. "Stochastic Process in Survival Analysis", volume-qin x bin, 14-December.

Zeger, Scott L., Diggle, Peter J., Liang, Kung-Yee, 2004. "A Cox Model For Biostatistics of the Future", Johns Hopkins University, Dep Biostatistics working papers, paper 32.