Original Article

Statistical methods for the analysis of increased access to emergency contraception

Irene Elizabeth Joy[1*], Anil C Mathew[2]

[1]Trainee Biostatistician, Department of Biostatistics, St. Thomas College, Pala, Arunapuram, INDIA.
[2]Professor of Biostatistics, Department of Community Medicine, PSG Institute of Medical Sciences and Research, Coimbatore, Tamil Nadu, INDIA.
Email: elaizairene910@gmail.com

Abstract

Survival analysis encompasses investigation of time to event data in clinical research. In survival analysis, it is common to have truncated data due to limited time span of the study or drop out of the subject for various reasons. The estimation of survivor function under left truncation was first discussed by Kaplan Meier by extending the well known product limit estimator of the survivor function. Later several parametric models are used for modelling survival data. The focus of this paper is to compare the models and their application in the study on increased access to emergency contraception by Raymond, E. *et al*. (2006) and their adverse impact on risk of sexual transmitted infections. Various models are computed using SAS (Statistical Analysis System) program. The findings are presented and discussed.
Keywords: emergency contraception.

[*]**Address for Correspondence:**
Dr. Irene Elizabeth Joy, Trainee Biostatistician, Department of Biostatistics, St. Thomas College, Pala, Arunapuram, INDIA.
Email: elaizairene910@gmail.com
Received Date: 08/11/2015 Revised Date: 19/12/2015 Accepted Date: 0/01/2016

Access this article online

Quick Response Code:

Website: www.statperson.com

DOI: 01 February 2016

INTRODUCTION

Survival Analysis is a set of methods and techniques used to analyse the times of transition among several states or conditions. The subject deals with modeling and analyzing data that correspond to the time from a well defined time origin until the occurrence of some particular event or end point[1,2]. Several statistical measures are developed to prepare a compact summary of the data that can explore their relationship between variables and time -to- event data in follow up study in clinical trial. One way to achieve this purpose is to search for a theoretical model (or distribution) that fits the observed data and identify the most important factors. The methods of survival analysis used in this paper are parametric and non parametric methods. We have presented various models in survival analysis: exponential, weibull and Kaplan Meier. The aim of this study is to develop appropriate programme on SAS (Statistical Analysis System) in various parametric and non- parametric model and to compare the outcome using a published data. The purpose of the presentation is to apply the models in the study and compare the results.

MATERIAL AND METHODS

NON PARAMETRIC METHOD

This method is free of assumptions. Furthermore, nonparametric analyses are more widely used in situations, where there is doubt about the exact form of distribution. In the nonparametric methods, the most popular and commonly used method is the Kaplan-Meier method. The Kaplan-Meier (KM) estimator is also called a non parametric maximum likelihood estimator. It is used for estimating the survival probabilities.

KAPLAN MEIER METHOD

Kaplan-Meier estimate is one of the best options to be used to measure the fraction of subjects living for a certain amount of time after treatment. In clinical trials or community trials, the effect of an intervention is assessed by measuring the number of subjects survived or saved after that intervention over a period of time. The Kaplan-Meier estimate is the simplest way of computing the survival over time in spite of all these difficulties associated with subjects or situations. The survival curve can be created assuming various situations. It involves computing of probabilities of occurrence of event at a certain point of time and multiplying these successive

How to site this article: Irene Elizabeth Joy, Anil C Mathew. Statistical methods for the analysis of increased access to emergency contraception. *International Journal of Recent Trends in Science and Technology* February 2016; 18(1): 13-16 http://www.statperson.com (accessed 02 February 2016).

probabilities by any earlier computed probabilities to get the final estimate. This can be calculated for two groups of subjects and also their statistical difference in the survivals.

$S(t)$ be the probability that a member from a given population will have a lifetime exceeding time, t. For a sample of size N from this population, let the observed times until death of the N sample members be

$$t_1 \leq t_2 \leq t_3 \leq \cdots \leq t_N.$$

The Kaplan–Meier estimator is the non parametric maximum likelihood estimate of $S(t)$, where the maximum is taken over the set of all piecewise constant survival curves with breakpoints at the event times t_i. It is a product of the form,

$$\hat{S}(t) = \prod_{t_i<t} \frac{n_i - d_i}{n_i}.$$

t_i : *survival time.*

n_i: the number "at risk" just prior to time t_i .

d_i: the number of deaths at time t_i.

When there is no censoring, n_i is just the number of survivors just prior to time t_i. With censoring, n_i is the number of survivors minus the number of losses (censored cases). It is only those surviving cases that are still being observed (have not yet been censored) that are "at risk" of an (observed) death.

PARAMETRIC METHOD

A parametric survival model is one in which survival time (the outcome) is assumed to follow some family of distributions. Exact distribution is unknown if parameters are unknown. It is used to estimate parameters.

ACCELERATED FAILURE TIME (AFT) MODEL

Accelerated failure time (AFT) models are alternatives to relative risk models which are used extensively to examine the covariate effects on event times in censored data regression. AFT models describe this "stretching out" or contraction of survival time as a function of predictor variables. Nevertheless, AFT models have been much less utilized in practice due to lack of reliable computing methods and software. In the statistical area of survival analysis, an accelerated failure time model (AFT model) is a parametric model that provides an alternative to the commonly used proportional hazard models. Whereas a proportional hazards model assumes that the effect of a covariate is to multiply the hazard by some constant, an AFT model assumes that the effect of a covariate is to accelerate or decelerate the life course of a disease by some constant. This is especially appealing in a technical context where the 'disease' is a result of some mechanical process with a known sequence of intermediary stages. The underlying assumption for AFT models is that the effect of covariates is multiplicative (proportional) with respect to survival time.

EXPONENTIAL MODEL

The density function of exponential distribution with parameter λ,

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x \leq 0 \end{cases}$$

Survival function: $S(t) = e^{-\lambda t}$

The exponential has constant hazard function, $h(t)=\lambda$.

WEIBULL MODEL

The most common distribution for parametric modeling of survival data is the Weibull distribution. The Weibull distribution has a desirable property in that if the AFT assumption holds then the PH assumption also holds.

Assuming T ~Weibull (λ, p) with probability density function

$f(t) = \lambda p t^{p-1} \exp(-\lambda t^p)$, where $p > 0$ and $\lambda > 0$,

the hazard function is given by

$h(t) = \lambda p t^{p-1}$

p is called shape parameter:

If $p > 1$ the hazard increases

If $p = 1$ the hazard is constant (exponential model)

If $p < 1$ the hazard decreases

FITTING THE MODELS IN STATISTICAL ANALYSIS SYSTEM (SAS)

Fitting these models often involves complicated and tedious computations and requires computer softwares like SAS.

PARAMETRIC ACCELERATED FAILURE TIME MODELS: THE LIFEREG PROCEDURE

The LIFEREG fits parametric models to failure time data that can be left-censored, right-censored, or interval-censored. The log of the survival time is modeled as a linear effect of covariates and a random disturbance term, the distribution of which includes the Weibull, log-normal, and log-logistic distributions.

NON PARAMETRIC METHODS

Kaplan-Meier (K-M) survival curves can be generated using SAS PROC LIFETEST.

LIFETEST computes the Kaplan-Meier estimate of a survivor function and provides the log-rank test to compare the underlying hazards of two or more samples of right-censored data. You can also use this procedure to study the association between the failure time and a number of concomitant variables.

DATA

The stimulated data is from a study conducted in Nevada and North Carolina between October 2002 and June 2005. Sexually active women, 14-24 years old, were randomly assigned to two methods of access to emergency contraceptive pills: increased access (two packages of pills dispensed in advance with unlimited resupply at no

charge) or standard access (pills dispensed when needed at usual charges). The target enrolment was 1490 women. Participants were followed for 1 year to assess incidence of sexually transmitted infections.

RESULT

INTERPRETATION OF NON PARAMETRIC METHOD: KAPLAN MEIER

The mean survival time and its standard error were underestimated because the largest observation was censored and the estimation was restricted to the largest event time. Here more than 50% of the observations and the largest observation are censored hence we cannot find the survival time. This is a disadvantage of Kaplan Meier method.

INTERPRETATION OF PARAMETRIC METHOD

Many parametric models are accelerated failure time models rather than PH models. SAS (9.1.3) only runs the AFT form of parametric models. The AFT assumption is applicable for a comparison of survival times. The following output and interpretation is based on the AFT assumption.

EXPONENTIAL MODEL

The exponential model assumes a constant hazard. This is indicated in the output by the value of the Weibull shape parameter (1.0000). The output can be used to calculate the estimated hazard for any subject given a pattern of covariates. The estimated hazard ratio of 0.5188 is obtained by exponentiating the estimated coefficient(-.6562) of the treat variable. i.e. Hazard ratio(treat=1 vs 0)=exp(-0.6562)=0.5188. The median time to occur sexually transmitted infection was decreased 49% in treated group. Since hazard ratio is less than 1, exposure benefits survival. This implies that Increased access to emergency contraception benefits survival.

WEIBULL MODEL

The estimated hazard ratio of 0.5195 is obtained by exponentiating the estimated coefficient (-0.7023/1.0727) of the treat variable. i.e. Hazard ratio(treat=1 vs 0)=exp(-β/scale)= exp(-0.7023/1.0727)=0.5195. The median time to occur sexually transmitted infection was decreased 49% in treated group. Since hazard ratio is less than 1, exposure benefits survival. This implies that Increased access to emergency contraception benefits survival.
i.e. Treatment is positive.

APPENDIX

SAS PROGRAM

```
proc import datafile="C:\Np\survire.sav" out=surv
dbms=sav replace;
run;
proc print data=surv;
run;
proc lifetest data=surv plots=(s);
time timepr*pregevt(0);
run;
proc lifereg data=surv;
model timepr*pregevt(0)=hemeth/dist=exponential;
run;
proc lifereg data=surv;
model timepr*pregevt(0)=hemeth/dist=weibull;
run;
```

OUTPUT OF KAPLAN MEIER

SUMMARY STATISTIC OF TIME VARIABLE

Table 1: Quartile estimate

Percent	Point estimate	95% Confidence Lower	Interval Upper
75	.	.	.
50	.	.	.
25	.	.	.

Mean: 344.354 **Standard error:** 1.674

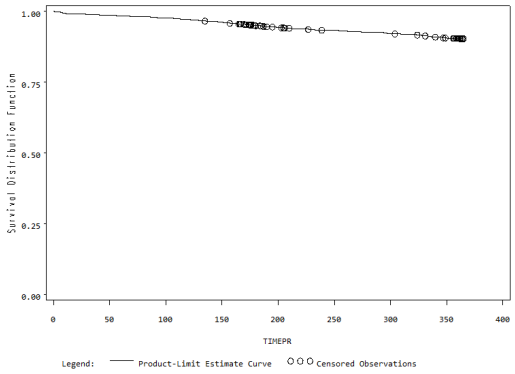

Figure 1(b): Kaplan meier graph

OUTPUT OF EXPONENTIAL

Table 2: Analysis of Parameter Estimates

Parameter	DF	Estimate	Standard Error	95%	Confidence Limits	Chi-Square	Pr > ChiSq
Intercept	**1**	7.9065	0.1054	7.6999	8.1131	5626.10 13.30	<.00010.0003
HEMETH	**1**	0.6562	0.1800	0.3035	1.0090		
Scale	**0**	1.0000	0.0000	1.0000	1.0000		
Weibull Shape	**0**	1.0000	0.0000	1.0000	1.0000		

OUTPUT OF WEIBULL

Table 3: Analysis of Parameter Estimates

Parameter	DF	Estimate	Standard Error	95%	Confidence Limits	Chi- Square	Pr > ChiSq
Intercept	**1**	8.0563	0.2178	7.6295	8.4832	1368. 12.17	<.0001 0.0005
HEMETH	**1**	0.7023	0.2013	0.3077	1.0970		
Scale	**1**	1.0727	0.0903	0.9096	1.2651		
Weibull Shape	**1**	0.9322	0.0785	0.7904	1.0994		

CONCLUSION

The enhanced access to emergency contraception had no adverse impact on risk of sexually transmitted infections (STI).

ACKNOWLEDGEMENTS

The authors are extremely thankful to Dr. S. Ramalingam, Principal, PSG Institute of Medical Sciences and Research for permitting us to do this study. We are grateful to Dr.Thomas V Chacko, Professor and Head of Department of Community Medicine, PSG Institute of Medical Sciences and Research, for his constant support and encouragement for the successful completion of the study. I am also thankful to Dr M Vijayakumar, Post Graduate Final Year, Department of Community Medicine, PSG Institute of Medical Sciences and Research, Coimbatore for his valuable comments which helped in improving the quality of this paper.

REFERENCE

1. D.R.Cox and D.Oakes, Analysis of Survival Data, Chapman and Hall London (1984).
2. Elandt-Johnson, R. C. and N.L. Johnson Survival Models and Data Analysis, John Wiley and Sons New York (1980).
3. D.Collett, Modeling Survival Data in Medical Research, Chapman and Hall London (1994).
4. J R. D W Hosmer, S.Lemeshow Applied Survival Analysis; Regression Modeling of Time to Event Data. John Wiley and Sons New York 1999.
5. D G Kleinbaum , Survival Analysis: A self-Learning Text. Springer-Verlag New York 1996.
6. P. D Allison, Survival Analysis Using the SAS System: A Practical Guide, SAS Institute, Cary, NC, SAS Institute, Inc., 1995.

Source of Support: None Declared
Conflict of Interest: None Declared