The First International Symposium on Optimization and Systems Biology (OSB'07)
Beijing, China, August 8–10, 2007

The Optimum Method of Prior Parameters in Survival Analysis*

Ming Han†

Department of Mathematics and Physics, Fujian University of Technology
Fuzhou, Fujian 350014, China

Abstract This paper introduces a new way for obtaining the optimum solution of prior parameters in survival analysis. Firstly, we give the definition of deviation limit, and discuss a property of prior parameters. Secondly, we present a method for obtaining the optimum solution of prior parameters with exponential distribution. Finally, we deal with a real data set, through which we can see that the presented method is not only efficient, but also easy to operate.

Keywords Survival analysis; hazard rate; optimum method; deviation limit; prior parameters.

1 Introduction

Consider type I censored life testings. Let the triple (n_i, r_i, t_i) be the data set of the ith testing, where the testings happen at times $t_1, t_2, \cdots, t_m$, n_i is the number of patients, r_i is the number of decedent among the patients.

If the survival time of the patient's life has the exponential distribution with parameter λ, i.e., it has the probability density function

$$f(t) = \lambda \exp\{-t\lambda\}, \quad t > 0, \tag{1}$$

where $\lambda > 0$, λ is the hazard rate of the exponential distribution (1).

For the exponential distribution (1), we take the conjugate prior of λ, Gamma(a, b), with density function

$$\pi(\lambda | a, b) = b^a \lambda^{a-1} \exp(-b\lambda)/\Gamma(a), \tag{2}$$

where $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$ is Gamma function, $a > 0$, $b > 0$, and both a and b are prior parameters.

*This work was supported partly by the Fujian Province Natural Science Foundation of China, and partly by the Fujian University of Technology.

†Email: hanming618@21cn.com

2 A Property of Prior Parameters and the Definition of Deviation Limit

2.1 A Property of Prior Parameters

For the exponential distribution (1), perform m times of type I censored life testings, denote the corresponding testing data set by (n_i, r_i, t_i), $i = 1, 2, \cdots, m$. Let $T = \sum_{i=1}^{m} (n_i - r_i) t_i, r = \sum_{i=1}^{m} r_i$. According to Lawless[1], the MLE (maximum likelihood estimator) of λ is

$$\widehat{\lambda}_C = \frac{r}{T}. \tag{3}$$

If we take the conjugate prior of λ, $Gamma(a, b)$, then, according to Berger [4], the Bayesian estimator of λ is

$$\widehat{\lambda}_B = \frac{a + r}{b + T}. \tag{4}$$

According to Kapur and Lamberson [3], MSE$(\widehat{\lambda}) = E(\widehat{\lambda} - \lambda_0)^2$, where λ_0 is the actual value of parameter λ, and $\widehat{\lambda}$ is an estimator of parameter λ. We generalize the results of Han[2], and deduce the following Theorem.

Theorem 1. For the testing data set (n_i, r_i, t_i) coming from m times of type I censored testings with life distribution (1) $(i = 1, 2, \cdots, m)$, let $T = \sum_{i=1}^{m} (n_i - r_i) t_i, r = \sum_{i=1}^{m} r_i$, suppose the prior density function of λ is given by (2),

(i) If MSE$\widehat{\lambda}_B$ <MSE$\widehat{\lambda}_C$, then a and b satisfy the inequality

$$(a - b\lambda_0)^2 - \frac{\lambda_0}{T}(b + T)^2 + \lambda_0 T < 0. \tag{5}$$

(ii) If $E(\widehat{\lambda}_B) = \lambda_0$, then a and b satisfy the equality

$$b\lambda_0 - a = 0. \tag{6}$$

Proof. **(i)** According to (3), (4) and the definition of mean square error (briefly MSE), the inequality MSE$\widehat{\lambda}_B$ <MSE$\widehat{\lambda}_C$ is equivalent to

$$E\left(\frac{a + r}{b + T} - \lambda_0\right)^2 - E\left(\frac{r}{T} - \lambda_0\right)^2 < 0. \tag{7}$$

According to Lawless[1], r obeys the Poisson distribution with parameter $\lambda_0 T$, so we have

$$E(r) = \lambda_0 T,\ E(r^2) = \lambda_0 T + \lambda_0^2 T^2. \tag{8}$$

According to (7) and (8), we get

$$E\left(\frac{a+r}{b+T}-\lambda_0\right)^2 - E\left(\frac{r}{T}-\lambda_0\right)^2$$
$$= \frac{1}{(b+T)^2}[(a-b\lambda_0)^2 - \frac{\lambda_0}{T}(b+T)^2 + \lambda_0 T]$$
$$< 0,$$

thus we have $(a-b\lambda_0)^2 - \frac{\lambda_0}{T}(b+T)^2 + \lambda_0 T < 0$.

(ii) If $E(\widehat{\lambda}_B) = \lambda_0$, that is $\lambda_0 = E(\widehat{\lambda}_B) = E\left(\frac{a+r}{b+T}\right) = \frac{a+\lambda_0 T}{b+T}$, then a and b satisfy $b\lambda_0 - a = 0$.

That concludes the proof of Theorem 1.

2.2 The Definition of Deviation Limit

According to (6) of Theorem 1, $\widehat{\lambda}_B$ is unbiased estimator of λ, which implies that the expectation of $Gamma(a,b)$ is λ_0, i.e. $E(\widehat{\lambda}_B) = \lambda_0$.

Definition 1. If $(a-b\lambda_0)^2 \leq M$, then the number M is called the deviation limit of expectation of $Gamma(a,b)$.

According to Definition 1, the validation of (5) is equivalent to that there exists a number M, such that a and b satisfy

$$\begin{cases} (a-b\lambda_0)^2 \leq M, \\ \frac{\lambda_0}{T}(b+T)^2 - \lambda_0 T > M. \end{cases} \tag{9}$$

3 The Optimum Solution of Prior Parameters

Generally, λ_0 is not exactly known, instead, λ_0 has an admissible range $[\lambda_L, \lambda_U]$. If M, λ_L and λ_U are all known, then from (9), we have

$$\begin{cases} b > \sqrt{MT/\lambda_0 + T^2} - T, \\ b\lambda_0 - \sqrt{M} \leq a \leq b\lambda_0 + \sqrt{M}. \end{cases} \tag{10}$$

When $\lambda_0 \in [\lambda_L, \lambda_U]$, we can obtain the admissible range of a and b , which is denoted by $R(\lambda_0)$ (see Figure 1). When $\lambda_0 = \lambda_L$ or $\lambda_0 = \lambda_U$, the corresponding range $R(\lambda_L)$ and $R(\lambda_U)$ are shown in Figure 2.

Based on the minimization principle of MSE (mean square error), optimum solution of a and b satisfy:

$$\begin{cases} min\{MSE(\widehat{\lambda}_B) - MSE(\widehat{\lambda}_C)\}, \\ a > 0,\ b > 0, \end{cases}$$

that is

$$\begin{cases} min\{(a-b\lambda_0)^2 - \frac{\lambda_0}{T}(b+T)^2 + \lambda_0 T\}, \\ a > 0,\ b > 0, \end{cases}$$

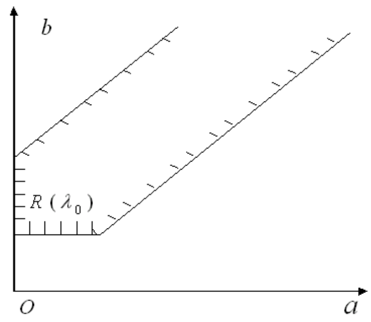

Figure 1: Range of a and b.

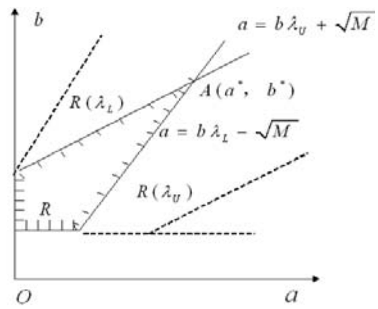

Figure 2: Optimum solution range of a and b.

or

$$\begin{cases} min(M-N) \\ (a-b\lambda_0)^2 = M, \\ \frac{\lambda_0}{T}(b+T)^2 - \lambda_0 T \geq N, \\ \lambda_L \leq \lambda_0 \leq \lambda_U, \\ a>0,\ b>0, \end{cases} \tag{11}$$

which is equivalent to (12)

$$\begin{cases} min(M-N) \\ (a-b\lambda_0)^2 = M, \\ \frac{\lambda_0}{T}(b+T)^2 - \lambda_0 T - c = N, \\ \lambda_L \leq \lambda_0 \leq \lambda_U, \\ a>0,\ b>0,\ c \geq 0. \end{cases} \tag{12}$$

If M, λ_L and λ_U are all given, then, from Figure 2, the optimum solution of a and b corresponds to the vertex $A(a^*, b^*)$ of convex set R while $c = 0$.

Notice that a^*, b^* satisfy:

$$\begin{cases} a = b\lambda_U - \sqrt{M}, \\ a = b\lambda_L + \sqrt{M}, \end{cases}$$

hence we have

$$\begin{cases} a = \frac{(\lambda_U+\lambda_L)\sqrt{M}}{\lambda_U-\lambda_L}, \\ b = \frac{2\sqrt{M}}{\lambda_U-\lambda_L}. \end{cases} \tag{13}$$

4 An Example

For a testing data set (n_i, r_i, t_i) coming from 3 times of type I censored testings with life distribution (1)($i = 1,2,3$), suppose that $t_1 = 20$(days), $t_2 = 50$(days), $t_3 = 100$(days), $n_1 = 100, n_2 = 200, n_3 = 200, r_1 = 1, r_2 = 2, r_3 = 3$, we have $T = \sum_{i=1}^{m}(n_i - r_i)t_i = 31580, r = \sum_{i=1}^{3} r_i = 6$. Furthermore, suppose that $\lambda_L = 1.5 \times 10^{-5}$, $\lambda_U = 2.5 \times 10^{-5}$, according to (3), we have $\widehat{\lambda}_C = \frac{r}{T} = 1.89993 \times 10^{-4}$.

From (13), we have

$$\begin{cases} a = \frac{(\lambda_U+\lambda_L)\sqrt{M}}{\lambda_U-\lambda_L} = 1.127659\sqrt{M}, \\ b = \frac{2\sqrt{M}}{\lambda_U-\lambda_L} = 8510.638298\sqrt{M}, \end{cases}$$

and from (4), we have

$$\widehat{\lambda}_B = \frac{a+r}{b+T} = \frac{1.127659\sqrt{M}+6}{8510.638298\sqrt{M}+31580} \tag{14}$$

Assume $M = 0.1, 0.2, 0.3, 0.4, 0.5$, according to (14), we can obtain $\widehat{\lambda}_B$, which is listed in Table 1.

Table 1 The Results of $\widehat{\lambda}_{iB}$

i	$\widehat{\lambda}_{iB}$	M	a	b
1	1.85478×10^{-4}	0.1	0.356597	2691.300136
2	1.83809×10^{-4}	0.2	0.504304	3806.073153
3	1.82598×10^{-4}	0.3	0.617644	4661.468575
4	1.81621×10^{-4}	0.4	0.713194	5382.600273
5	1.80791×10^{-4}	0.5	0.797375	6017.930053

From Table 1, we find that for different $M(M = 0.1, 0.2, 0.3, 0.4, 0.5)$, $\widehat{\lambda}_{iB}$ are all robust, and $\widehat{\lambda}_{iB} < \widehat{\lambda}_C$. Based on Table 1, we can obtain the estimates $\widehat{S}_i(t)$ of the survival function at the time t, which is listed in Table 2.

Table 2 The Results of $\widehat{S}_i(t)$

$\widehat{S}_i(t)$	M	100	200	300	400	500
$\widehat{S}_1(t)$	0.1	0.9816232	0.9635840	0.9458764	0.9284941	0.9114314
$\widehat{S}_2(t)$	0.2	0.9817870	0.9639057	0.9463501	0.9291142	0.9121923
$\widehat{S}_3(t)$	0.3	0.9819059	0.9641392	0.9466940	0.9295644	0.9127448
$\widehat{S}_4(t)$	0.4	0.9820018	0.9643276	0.9469715	0.9299277	0.9131907
$\widehat{S}_5(t)$	0.5	0.9820833	0.9643276	0.9472073	0.9302365	0.9135698
$\widehat{S}_-$		0.0004601	0.0007436	0.0013309	0.0017424	0.0021384

Note: $\widehat{S}_i(t) = \exp(-\widehat{\lambda}_{iB}t)$, in the Table 2, and $\widehat{\lambda}_{iB}$ are given by Table 1($i = 1,2,3,4,5$), which is the corresponding $M = 0.1, 0.2, 0.3, 0.4, 0.5$, $\widehat{S}_- = \widehat{S}_5(t) - \widehat{S}_1(t)$.

From Table 2, we find that for different M($M = 0.1, 0.2, 0.3, 0.4, 0.5$), $\widehat{S}_i(t)$ are all robust($i = 1,2,3,4,5$).

5 Conclusion

From (13), we find that if $M = 0$, then $a = b = 0$, and then $\widehat{\lambda}_B = \widehat{\lambda}_C$. For Gamma distribution—— Gamma (a,b), where $a > 0, b > 0$, we know that $M > 0$, $\widehat{\lambda}_B < \widehat{\lambda}_C$.

From the method of obtaining optimum solution for prior parameters, if a and b satisfy MSE$\widehat{\lambda}_B$ <MSE$\widehat{\lambda}_C$, then $\widehat{\lambda}_B$ is superior to $\widehat{\lambda}_C$ according to minimization principle of MSE.

Acknowledgments

The author has benefited from Professor Quan-Ling Wei, who has checked the paper and given author very helpful suggestions.

References

[1] Lawless J.F. (1982) *Statistics models and methods for lifetime data*. New York, John Wiley & Sons.

[2] Han M.: Suitable regions of prior distribution exponential distribution of fixed-time gross-ending. (1997) *Chinese Mathematical Statistics and Applied Probability*. 12(1): 81-85.

[3] Kapur K.C. and Lamberson L.R. (1977) *Reliability in engineering design*. New York, John Wiley & Sons.

[4] Berger J.O. (1985) *Statistical Decision Theory and Bayesian Analysis*. Second edition, New York, Springer-Verlag.