استخدام طرق متعددة لتقدير معلمات احتمالية دالة كثافة الوفيات

مع تطبيق عملي

Using Different Methods to Estimate the Parameters of Probability Death Density Function with Application

أ.م.د.أسيل سمير محمد

قسم طب المجتمع / كلية طب الكندي

المستخلص

يتناول هذا البحث اشتقاق مقدرات لتقدير معلمة القياس (λ) ومعلمة الشكل (γ) لأحدى التوزيعات الاحتمالية للفشل للمرضى الذين يعانون من مرض الصداع، وبعد تقدير المعلمات يقدر متوسط بقاء المريض وهو مرتاح من هذا الصداع، وهذا مهم جداً للأطباء والباحثين في حقول التجارب الطبية والحياتية. وبعد استخراج المقدرات بأربعة طرائق هي الإمكان الأعظم، ومقدرات وايت (White)، ومقدر بيز (Bayes) ومقدرات (Bain and Antle)، طبقت النتائج على بيانات حقيقية وهذا ما سعى اليه الباحث لتوظيف التطبيق الإحصائي في المجالات الصحية.

Abstract

In this paper, we derive a formula for estimating the two parameters (λ, γ) for two parameter's Weibull, where (λ) is scale parameter and (γ) is shape parameter using the method of maximum likelihood and White estimators and Bayesian estimator, where (λ) is considered as random variable having prior distribution $[g(\lambda)]$, the Bayes estimator of (λ) is derived according to squared error function, and then the estimators are applied to the set of data represent relief time (in hours) for (20) patient suffering from headache pain, we found that White method is simplest method.

Key words maximum likelihood estimator, white estimators, Bayes estimator, γ is shape parameter, λ is scale parameter for two parameters Weibull distribution.

$1 - Aim\ of\ the\ Research$

The aim of this research is to derive an estimator for the parameters of probability death density function for remission times of (20) patients suffering from a headache disease. The distribution studied is two parameters Weibull (λ, γ) where (λ) is scale parameter and (γ) is shape parameter. The method are maximum

likelihood, White and Bayesian estimators. Finally Bain and Antle estimator, the estimated parameters are used to estimate the mean relief time, which is necessary for doctors working on medical experiment to improve the treatement.

$2 - Introduction$

Let $t_1, t_2, \ldots, t_n$ are independent survival times for n individuals, each of them have a constant hazard rate λ. Let $y = \sum_{i=1}^{n} t_i$ i.e, y is total survival time of all n individuals, the probability death density function can be determind using mathematical induction procedure, many probability death density can be demonstrated through applying mathematical induction on hazard rate function. Some application of the hazard rate is the studying a group of patients suffering from a kidney disease, in which the failure rate of each kidney is constant and equals, for a patient to die from this disease. Both kidneys must fail. Hence the hazard rate of these patients with this disease is $\lambda[\lambda t|(1 + \lambda t)]$ and it is increasing function of tie for $\lambda > 0$.

We consider here a widely used hazard rate which is the Weibull hazard rate given in the equation:

$$\lambda(t) = \lambda\gamma(t - \delta)^{\gamma - 1} \qquad (1)$$

λ, γ, δ are real parameters and > 0

If $t = (t - \delta)$ then we can write Weibull hazard rate $\lambda(t)$ as:

$$\lambda(t) = \lambda\gamma t^{\gamma - 1} \qquad (2)$$

Then the death density function $f(t)$, and survival distribution of $y = \sum_{i=1}^{n} t_i$ are:

$$f(t) = \lambda\gamma t^{\gamma - 1} e^{-\lambda t^{\gamma}} \qquad (3)$$

$$S(t) = e^{-\lambda t^{\gamma}} \qquad (4)$$

And cumulative distribution function

$$C.D.F = F(t) = 1 - S(t)$$

$$F(t) = 1 - e^{-\lambda t^{\gamma}} \qquad (5)$$

This paper is organized as follows, section (1) include introduction and section (2) deals with introducing different methods to estimate the shape parameter (γ) and the scale parameter (λ), with simulation procedure to obtain the best estimator which have smallest mean square error (MSE), while section (3) include the application of the above estimators, on the data for 20 patients.

$3 - Estimation\ Method$

$3 - 1\ maximum\ likelihood\ estimator$

let $T_1, T_2, \dots, T_n$ be a random sample from $f(t)$ defined in (3), then the joint function;

$$L = \prod_{i=1}^{n} f(t_i; \lambda, \gamma) = \lambda^n \gamma^n \pi\, t_i^{\gamma-1} e^{-\lambda \sum t_i^{\gamma}} \qquad (6)$$

$$\ln L = n \ln \lambda + n \ln \gamma + (\gamma - 1) \sum \ln t_i - \lambda \sum t_i^{\gamma} \qquad (7)$$

Differentiating $\ln L$ with respect to λ, γ;

$$\frac{\partial \ln L}{\partial \lambda} = \frac{n}{\lambda} - \sum t_i^{\gamma} \qquad (8)$$

$$\frac{\partial \ln L}{\partial \gamma} = \frac{n}{\gamma} \qquad \log t_i - \lambda \sum t_i^{\gamma} (1) \log t_i \qquad (9)$$

From (8) and (9) we get;

$$\frac{n}{\hat{\lambda}} = \sum t_i^{\hat{\gamma}}$$

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} t_i^{\hat{\gamma}}} \qquad (10)$$

which is an implicit function of $\hat{\gamma}$

when $\hat{\gamma} = 1 \Rightarrow \hat{\lambda}_{MLE} = \frac{1}{\bar{t}}$ and from (9);

$$\frac{n}{\hat{\gamma}} + \sum \ln t_i - \hat{\lambda} \sum t_i^{\hat{\gamma}} (1) \log t_i = 0$$

$$\sum \ln t_i - \hat{\lambda}_{MLE} \sum (\log t_i)\, t_i^{\hat{\gamma}} = -\frac{n}{\hat{\gamma}}$$

$$\hat{\gamma}_{MLE} = \frac{n}{\hat{\lambda}_{MLE} \sum (\log t_i) t_i^{\hat{\gamma}} - \sum \ln t_i} \qquad (11)$$

$3-2$ *White Estimators*

This method depend on simple linear regression to estimate (λ, γ), using F(t).

$$F(t_i) = 1 - e^{-\lambda \sum t_i^{\gamma}}$$

$$e^{-\lambda \sum t_i^{\gamma}} = 1 - F(t_i)$$

$$\lambda t_i^{\gamma} = -[\log(1 - F(t_i))]$$

$$\log \lambda + \gamma \log t_i = \log[-\log(1 - F(t_i))]$$

Comparing with;

$$E(y_i) = \alpha + \beta E(x_i)$$

$$y_i = \log \lambda + \gamma (\log t_i)$$

where, $x_i = \log t_i$, $\alpha = \log \lambda$, $\beta = \gamma$

Since the least square estimator of β *is* $\hat{\beta}$ *and of* α *is* $\hat{\alpha}$

$$\hat{\beta} = \frac{\sum x_i y_i - n\bar{x}\,\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

Therefore; $\hat{\alpha} = \log \hat{\lambda}$

$$\hat{\lambda} = e^{\hat{\alpha}} \qquad (12)$$

$$\hat{\beta} = \hat{\gamma} \qquad (13)$$

$3-3$ *Bayesian Estimators*

Suppose the scale parameter λ is random variable having prior distribution $g(\lambda)$,

$$g(\lambda) = K\frac{e^{-a\lambda}}{\lambda^{c}} \qquad 0 < \lambda < \infty \qquad (14)$$

Then $g(\lambda|t)$ is the posterior distribution of λ given t is found from;

$$g(\lambda|t) = \frac{\pi f(t_i;\lambda,\gamma)g(\lambda)}{\int_0^\infty f(t_i;\lambda,\gamma)g(\lambda)d\lambda} \qquad (15)$$

$$f(t) = \int_0^\infty \lambda^n \gamma^n \pi\ t_i^{\gamma-1}\ e^{-\lambda\sum t_i^\gamma}\ \frac{e^{-a\lambda}}{\lambda^c}\ d\lambda$$

$$f(t) = \gamma^n \pi\ t_i^{\gamma-1} \int_0^\infty \lambda^{n-c}\quad e^{-\lambda(\sum t_i^\gamma + a)}\ d\lambda$$

$$= \frac{\gamma^n \pi\ t_i^{\gamma-1}}{(\sum t_i^\gamma + a)^{n-c+1}}\ \Gamma(n-c+1)$$

$$g(\lambda|t) = \frac{\gamma^n \pi\ t_i^{\gamma-1}\lambda^{n-c}e^{-\lambda(\sum t_i^\gamma + a)}}{\frac{\gamma^n \pi\ t_i^{\gamma-1}\Gamma(n-c+1)}{(\sum t_i^\gamma + a)^{n-c+1}}}$$

$$= \frac{(\sum t_i^\gamma + a)^{n-c+1}}{\Gamma(n-c+1)}\lambda^{n-c}e^{-\lambda(\sum t_i^\gamma + a)} \qquad (16)$$

which represents Gamma distribution with $(n-c+1, \sum t_i^\gamma + a)$

Then according to the squared error function $(\lambda - d)^2$, the Bayes estimator of parameter λ (when γ is known) is giving by;

$$\hat{\lambda}_{Bayes} = \frac{\sum t_i^\gamma + a}{n-c+1} \qquad (17)$$

where a, c are constant, and (γ is known) or estimated.

$3-4$ *Bain and Antle Estimators*

The fourth method for estimating the parameters of common death density function Weibull (λ,γ) is called (Bain & Antle) estimators, where the estimators (λ^*,γ^*) are suggested by (Bain & *Antle* 1967), which jointly maximize the agreement between $u_{(i)}$ and $E(u_{(i)}), u_{(i)}$ is the ith largest value of $u(t_i;\lambda,\gamma)$ in

the sample. We apply the least squares method to maximize this agreement, thus (λ^*& γ^*) are chosen so that;

$$T = \sum_{i=1}^{n}\left[\tau\left(u_{(i)}\right) - \tau(E\left(u_{(i)}\right))\right]^2 \tag{18}$$

τ is monotone function of both $u_{(i)}$ and $E\left(u_{(i)}\right)$, let;

$$u(t_i; \lambda, \gamma) = \lambda\, t^{\gamma} \tag{19}$$

This choice of u gives the density function;

$$f(u) = e^{-u}$$

Let $u_{(1)} < u_{(2)} < \cdots < u_{(n)}$ be ordered values of sample, then;

$$E\left(u_{(i)}\right) = z_i = \sum_{j=1}^{i} \frac{1}{n-j+1} \tag{20}$$

To obtain closed estimators (λ^*& δ^*), we chose τ as the natural logarithm function , then minimize the function (21);

$$\tau_1(\lambda, \gamma) = \sum_{i=1}^{n}\left(log\, \lambda\, t_{(i)}^{\gamma} - log\, z_i\right)^2 \tag{21}$$

with respect to λ and γ , this yield the estimators;

$$\lambda^* = \left[\frac{\prod_{i=1}^{n} z_i}{\prod_{i=1}^{n} t_{(i)}^{\gamma^*}}\right]^{\frac{1}{n}} \tag{22}$$

And,

$$\gamma^* = \frac{\sum_{i=1}^{n} \log z_i \log t_{(i)} - \frac{\sum_{i=1}^{n} \log z_i \sum_{i=1}^{n} \log t_{(i)}}{n}}{\sum_{i=1}^{n} (\log t_{(i)})^2 - \frac{[\sum_{i=1}^{n} \log t_{(i)}]^2}{n}} \tag{23}$$

$4 - Application$

The following data represents recorded relief time (in hours) for 20 patients taking an analgesic to relive headache pain:

t_i	1.1	1.4	1.3	1.7	1.9	1.8	1.6	2.2	1.7	2.7
	4.1	1.8	1.5	1.2	1.4	3.0	1.7	2.3	1.6	2.0

Starting with $\hat{\gamma}_0 = 0.5$ solving equation (9), we find the maximum likelihood estimator of $\hat{\gamma}_{MLE} = 2.79$ applying this in (10), we find $\hat{\lambda}_{MLE} = 0.12$, according to this $\hat{\gamma} = 2.79$ & $\hat{\lambda} = 0.12$.

The estimating mean $(\hat{\mu})$ (which represent the mean relief time);

$$\hat{\mu} = \hat{\lambda}^{-1/\hat{\gamma}}\, \Gamma(1 + \hat{\gamma}) = 1.89 \ hours$$

From equations (21) & (22), the Bain & Antle estimators of;

$$\hat{\lambda}^* = 0.081 \qquad \hat{\gamma}^* = 3.44 \qquad \mu^* = 1.87 \ hours$$

Bain & Antle estimators are unbiased in small samples, it is less bias than maximum likelihood estimators in small samples.

However, for large sample properties of the maximum likelihood estimators still make them quite attractive for sample sizes of 50 or more. Also the white estimators of λ, γ which are denoted here $\lambda^{**}, \gamma^{**}$ are found from the data, and equation (12), $\hat{\lambda}^{**} = 0.085$, equation (13), we find $\hat{\gamma}^{**} = 3.35$, also $\mu^* = 1.87 \ hours$. The following table summarize the results of estimated parameters from data.

Estimates	λ	γ	μ
$\hat{\lambda}, \hat{\gamma}$	0.12	2.79	1.89
$\hat{\lambda}^*, \hat{\gamma}^*$	0.081	3.44	1.87
$\hat{\lambda}^{**}, \hat{\gamma}^{**}$	0.085	3.35	1.89
$\hat{\lambda}_{Bayes}, \gamma_{know}$	0.092	3	1.88

For all estimators the estimated $(\hat{\mu})$ (mean relief time) $(\hat{\mu} \cong 1.88\ hours)$. This is considered as the mean time after it, the patient must take the treatment.

$5 -$ ***Conclusions***

1- The maximum likelihood estimators cannot be obtained in a closed form (equation 10 & 11), need numerical method to find the estimators of MLE, but this difficulties are not important due to the properties of MLE estimators.

2- In Bayes estimator, only the scale parameter (λ) is considered random variable having prior distribution $g(\lambda)$, while the shape parameter (γ) is considered known.

3- For Bain & Antle estimators, only we consider $u(t;\lambda,\gamma) = \lambda t^{\gamma}$ in order to obtain the simple estimator (λ^{*} & γ^{*}), we can use another form like $u(t;\lambda,\gamma) = (1 - \mathrm{e}^{-\lambda t^{\gamma}})$, but due to the nature of research which insisted on application of death density function, the selection $u(t;\lambda,\gamma) = \lambda t^{\gamma}$ is enough.

4- The estimator's by white method depend on simple regression procedure, which is depend on minimizing the sum squares of the differences between observed and expected values. The estimators of $\lambda\ is\ \lambda^{**}$ and of δ is δ^{**}.

5- All the estimators are obtained for complete sample size, it can be modified to the case of censored failure times.

6- All the estimators agree very well with $\hat{\mu} = 1.89\ hours$ (estimated mean relief time).

7- The concept of this research can be expanded to compute the variance of estimated parameters and to build confidence limits for estimators.

References

1- D. R. Thomas and W. M. Wilson (1972), " linear order statistic estimation for the two – parameter Weibull and extreme value distributions from type II progressively", considered examples, Technometrics, 14, 679 – 691.

2- T.R. Tsui, Y.- Lu and S. – J. Wu, (2008), " reliability sampling plan for weibull distribution with limited capacity of test facility" Computers and Industrial Engineering, 55, 721 – 728.

3- Giovana O.Silva .Edwin M. M. Ortega (2010) , " The beta modified Weibull distribution " Lifetimedata Anel(2010) 16:409-430

4- Roy,M.K,Podder ,C.K and Bhuiyan ,K.J.(2002),minimax estimation of the scale parameters of the weibull distribution for quadratic and mlinex loss functions jahangirnagar university journal of science , 25,277-285

5- Dey,s.(2008) minimax estimation of the parameter of the rayleigh distribution under quadratic loss function . data science journal, vol .7,pp,23-30.