

在定数截尾样本下三参数威布尔分布的矩估计 *

孙丽芬

费鹤良

(徐州师范大学, 徐州, 221116) (上海师范大学, 上海, 200234)

摘 要

本文讨论了在定数截尾样本下三参数威布尔分布的矩估计问题. 在定数截尾情形下, 将威布尔分布数据转化为均匀分布数据, 利用平均剩余寿命构造样本矩, 同时, 第三阶矩方程用样本的第一个次序统计量来代替, 得到了在定数截尾样本下三参数威布尔分布的矩估计方程, 用随机模拟方法得出了矩估计的偏性和均方误差. 并与近似 MLE 进行了比较, 表明此矩估计方法有较好的性质.

关键词: 三参数威布尔分布, 矩估计, 平均剩余寿命, 次序统计量.

学科分类号: O213.2.

§1. 引 言

三参数威布尔分布是可靠性中非常重要的寿命分布之一. 在定数截尾样本下三参数威布尔分布的极大似然估计求解十分困难, 且仅适用于形状参数在限定的范围之内. 陈迪 [2] 给出了一种近似极大似然估计 (AMLE), 可减少计算的困难. Cohen and Whitten[3], Cohen, Whitten and Ding Y.[4] 等讨论了在完全样本下三参数威布尔分布的矩估计. 对于不完全样本, 至今未见这方面的文献. 对于截尾样本的矩估计, 倪中新 [5] 给出了指数分布的矩估计.

本文利用平均剩余寿命构造样本矩, 并将复杂的威布尔分布样本转化为均匀分布样本进行讨论, 其中第三阶矩的方程用样本的第一个次序统计量来代替, 利用次序统计量的性质得出了矩方程, 从而得出了修正矩估计 (MME). 随机模拟结果显示, MME 较 AMLE 优越.

§2. 基本知识

假设 X 表示某产品的寿命, 其分布函数为 $F(X)$, 即 $X \sim F(X)$.

若产品工作到时刻 t 仍能正常工作, 能继续工作的时间为 X 的概率为 $F_t(x)$. 则

$$F_t(x) = P(X - t \leq x | X > t) = \begin{cases} \frac{F(x+t) - F(t)}{1 - F(t)}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

容易验证 $F_t(x)$ 为分布函数. 称之为剩余寿命分布函数.

令 $m(t)$ 为产品工作到时刻 t 仍能正常工作的条件下, 继续工作的平均时间. 则

$$m(t) = \int_0^{\infty} x dF_t(x) = \int_0^{\infty} (1 - F_t(x)) dx = \frac{1}{1 - F(t)} \left[E(X) - \int_0^t (1 - F(y)) dy \right],$$

称 $m(t)$ 为产品工作到时间 t 仍能正常工作时的平均剩余寿命.

* 国家自然科学基金资助项目 (19671058).

本文 2001 年 11 月 7 日收到, 2002 年 4 月 10 日收到修改稿.

若 $X \sim U(0, 1)$, 即 x 有密度函数

$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1; \\ 0, & \text{其它.} \end{cases}$$

易知

$$m(t) = \frac{1}{2}(1-t), \quad 0 \leq t \leq 1. \quad (1)$$

§ 3. 构造矩方程

假设某产品的寿命 $X \sim \text{Wei}(x; \beta, \delta, \gamma)$, 即

$$F(x) = \begin{cases} 1 - \exp\left[-\left(\frac{x-\gamma}{\beta}\right)^\delta\right], & x \geq \gamma; \\ 0, & x < \gamma, \end{cases}$$

其中 $\beta > 0$ 为尺度参数, γ 为门限或位置参数, $\delta > 0$ 为形状参数.

将 N 个产品投入试验, 到恰有 n ($n \leq N$) 个失效时停止试验. 特别地, 当 $n = N$, 即完全样本时, 已进行了充分的讨论. 对定数截尾情形, 即 $n < N$, 所获样本为 $X_{1:N} \leq X_{2:N} \leq \dots \leq X_{n:N}$. 若令

$$Y = 1 - \exp\left[-\left(\frac{X-\gamma}{\beta}\right)^\delta\right], \quad \text{其中 } a^+ \text{ 为 } a \text{ 的正部}, \quad (2)$$

即 $Y = F(X)$, 则 $Y \sim U(0, 1)$. 其期望与方差分别为 $E(Y) = 2^{-1}$, $\text{Var}(Y) = 12^{-1}$. 相应的次序统计量为 $Y_{1:N} \leq Y_{2:N} \leq \dots \leq Y_{n:N}$. 令

$$U = \frac{1}{N} \left[\sum_{i=1}^n Y_{i:N} + (N-n)(Y_{n:N} + m_Y(Y_{n:N})) \right], \quad (3)$$

$$S^2 = \frac{1}{N-1} \left\{ \left[\sum_{i=1}^n Y_{i:N}^2 + (N-n)(Y_{n:N} + m_Y(Y_{n:N}))^2 \right] - \frac{1}{N} \left[\sum_{i=1}^n Y_{i:N} + (N-n)(Y_{n:N} + m_Y(Y_{n:N})) \right]^2 \right\}, \quad (4)$$

可得以下结论.

定理 若 $Y \sim U(0, 1)$, 则

$$\begin{aligned} E(U) &= E(Y), \\ E(S^2) &= \omega \cdot \text{Var}(Y), \end{aligned} \quad (5)$$

其中, U, S^2 的定义分别同 (3), (4) 式,

$$\omega = \frac{n(n+1)(n+2) + 3(Nn-n^2)(N+2)}{N(N+1)(N+2)}.$$

证明: 由 [7] 知, 均匀分布的次序统计量有如下性质:

$$\begin{aligned} E(Y_{i:N}) &= \frac{i}{N+1}, & 1 \leq i \leq N, \\ E(Y_{i:N}^2) &= \frac{i(i+1)}{(N+1)(N+2)}, & 1 \leq i \leq N, \\ E(Y_{i,j:N}) &= \frac{i(j+1)}{(N+1)(N+2)}, & 1 \leq i < j \leq N, \\ \text{Var}(Y_{i:N}) &= \frac{i(N+1-i)}{(N+1)^2(N+2)}, & 1 \leq i \leq N. \end{aligned} \quad (6)$$

由 (1) 式知，对定数截尾情形，平均剩余寿命 $m_Y(Y_{n:N}) = 2^{-1}(1 - Y_{n:N})$.

$$\begin{aligned} E(U) &= E\left\{\frac{1}{N}\left[\sum_{i=1}^n Y_{i:N} + (N-n)(Y_{n:N} + m_Y(Y_{n:N}))\right]\right\} \\ &= \frac{1}{N}\left[\sum_{i=1}^n E(Y_{i:N}) + \frac{1}{2}(N-n)(1 + E(Y_{n:N}))\right]. \end{aligned}$$

将 (6) 式代入即得 $E(U) = 2^{-1} = E(Y)$.

$$\begin{aligned} E(S^2) &= \frac{1}{N-1}\left\{\left[\sum_{i=1}^n E(Y_{i:N}^2) + (N-n)E(Y_{n:N} + m_Y(Y_{n:N}))^2\right] \right. \\ &\quad \left. - \frac{1}{N}E\left[\sum_{i=1}^n Y_{i:N} + (N-n)(Y_{n:N} + m_Y(Y_{n:N}))\right]^2\right\} \\ &= \frac{1}{N}\sum_{i=1}^n E(Y_{i:N}^2) + \frac{nN-n^2}{4N(N-1)}E(Y_{n:N}^2) - \frac{2}{N(N-1)}\sum_{i=1}^{n-1}\sum_{j=i+1}^n E(Y_{i,j:N}) \\ &\quad + \frac{nN-n^2}{4N(N-1)} - \frac{N-n}{N(N-1)}\sum_{i=1}^n E(Y_{i,n:N}) - \frac{N-n}{N(N-1)}\sum_{i=1}^n E(Y_{i:N}) \\ &\quad + \frac{nN-n^2}{2N(N-1)}E(Y_{n:N}). \end{aligned}$$

将 (6) 式代入即得 $E(S^2) = \omega/12 = \omega \cdot \text{Var}(Y)$. #

根据上述定理，并引入第一个次序统计量 $Y_{1:N}$. 对均匀分布，我们让

$$\begin{aligned} E(Y) &= U, \\ \omega \cdot \text{Var}(Y) &= S^2, \\ E(Y_{1:N}) &= Y_{1:N}. \end{aligned} \tag{7}$$

将 (2), (3), (4) 式代入，又考虑到在实际问题中可令 $X_{1:N} \geq \gamma$ ，即得三参数威布尔的矩估计方程

$$\frac{1}{2} = \frac{1}{N}\left\{\sum_{i=1}^n \left[1 - \exp\left(-\left(\frac{X_{i:N} - \gamma}{\beta}\right)^\delta\right)\right] + \frac{N-n}{2}\left[1 - \exp\left(-\left(\frac{X_{n:N} - \gamma}{\beta}\right)^\delta\right) + 1\right]\right\}, \tag{8}$$

$$\begin{aligned} \frac{\omega}{12} &= \frac{1}{N-1}\left\{\left[\sum_{i=1}^n \left[1 - \exp\left(-\left(\frac{X_{i:N} - \gamma}{\beta}\right)^\delta\right)\right]^2 + \frac{N-n}{4}\left[1 - \exp\left(-\left(\frac{X_{n:N} - \gamma}{\beta}\right)^\delta\right) + 1\right]^2\right] \right. \\ &\quad \left. - \frac{1}{N}\left[\sum_{i=1}^n \left[1 - \exp\left(-\left(\frac{X_{i:N} - \gamma}{\beta}\right)^\delta\right)\right] + \frac{N-n}{2}\left[1 - \exp\left(-\left(\frac{X_{n:N} - \gamma}{\beta}\right)^\delta\right) + 1\right]\right]^2\right\}, \end{aligned} \tag{9}$$

$$\frac{1}{N+1} = 1 - \exp\left[-\left(\frac{X_{1:N} - \gamma}{\beta}\right)^\delta\right]. \tag{10}$$

由 (10) 式得：

$$\beta^\delta = -\frac{(X_{1:N} - \gamma)^\delta}{\ln[N/(N+1)]},$$

代入 (8), (9) 式并化简得：

$$\frac{1}{2} = \frac{1}{N}\left\{\sum_{i=1}^n \left[1 - \left(\frac{N}{N+1}\right)^{\left(\frac{X_{i:N}-\gamma}{X_{1:N}-\gamma}\right)^\delta}\right] + \frac{N-n}{2}\left[2 - \left(\frac{N}{N+1}\right)^{\left(\frac{X_{n:N}-\gamma}{X_{1:N}-\gamma}\right)^\delta}\right]\right\} \tag{11}$$

$$\frac{\omega}{12} = \frac{1}{N-1}\left\{\sum_{i=1}^n \left[1 - \left(\frac{N}{N+1}\right)^{\left(\frac{X_{i:N}-\gamma}{X_{1:N}-\gamma}\right)^\delta}\right]^2 + \frac{N-n}{4}\left[2 - \left(\frac{N}{N+1}\right)^{\left(\frac{X_{n:N}-\gamma}{X_{1:N}-\gamma}\right)^\delta}\right]^2 - \frac{N}{4}\right\}. \tag{12}$$

可利用迭代方法由第 (11), (12) 式得出 γ, δ 的矩估计，再由式 (10) 得出 β 的估计。

迭代过程中, 首先选取初值 γ_0, δ_0 , 再利用牛顿迭代法, 首先考虑一个方程, 再考虑另一个方程, 反复迭代, 直至两个方程的值都小于给定的数值为止. 其中对于每一组具体的数据, 初值选取要求 $\gamma_0 < x_{1:N}$. 这种限制在理论上是合理的, 且易行. δ_0 一般不会太大. 通过对同一组值, 取不同的初值进行迭代, 结果相差很小. 说明上述迭代的收敛性质是稳定的. 作者对数万组数据进行了模拟, 结果表明方程 (11)-(12) 的解存在且唯一.

§4. Monte-Carlo 模拟比较

我们通过 1000 次 Monte-Carlo 模拟, 得到了 MME 的相对偏差 $E[(\hat{\gamma} - \gamma)/\gamma]$, $E[(\hat{\beta} - \beta)/\beta]$, $E[(\hat{\delta} - \delta)]$, 与相对均方误差 $E[(\hat{\gamma} - \gamma)/\gamma]^2$, $E[(\hat{\beta} - \beta)/\beta]^2$, $E[(\hat{\delta} - \delta)]^2$ 的估计值.

由表 1, 表 2 易看出与 [2] 中提供的 AMLE 的相应数据相比, MME 的偏差, 均方误差都要明显的比 AMLE 的相应值小, 且截尾情况对 MME 的影响不大. 从实际结果看, MME 具有较好的性质.

表 1 $N = 30$ 时估计量的偏差

n	δ	BIAS					
		δ		γ		β	
		MME	AMLE	MME	AMLE	MME	AMLE
15	1	1.1838	3.9180	-0.1651	-0.8539	0.0066	0.2757
	2	0.1117	1.9808	-0.0370	-0.4020	0.0485	0.3618
	3	-0.6948	0.9848	0.0772	-0.1428	-0.0312	0.1490
	4	-0.5590	0.2941	0.0470	-0.0068	-0.0188	0.0295
22	1	1.1206	2.1538	-0.1986	-0.9037	0.2019	1.0103
	2	0.1089	0.3979	-0.0394	-0.0897	0.0488	0.0834
	3	-0.6921	0.0061	0.0760	0.0345	-0.0583	-0.0442
	4	-0.5610	-0.4758	0.0504	0.1160	-0.0400	-0.1248
30	1	1.2039	1.5685	-0.2306	-0.8813	0.2883	1.2292
	2	0.0940	0.1671	-0.0450	-0.0126	0.0565	0.0091
	3	-0.6464	-0.1195	0.0808	0.0966	-0.0837	-0.1169
	4	-0.5508	-0.5437	0.0447	0.1655	-0.0446	-0.1837

表 2 $N = 30$ 时估计量的均方误差

n	δ	MSE					
		δ		γ		β	
		MME	AMLE	MME	AMLE	MME	AMLE
15	1	1.5436	8.2167	0.0285	0.8171	0.0521	0.2390
	2	0.1889	9.6610	0.0057	0.4334	0.0323	0.3894
	3	0.7287	9.6846	0.0117	0.2793	0.0263	0.2576
	4	1.4212	13.0288	0.0075	0.2569	0.0169	0.2345
22	1	1.3925	5.3104	0.0433	0.9038	0.0853	1.3793
	2	0.1539	1.7518	0.0057	0.1508	0.0166	0.1856
	3	0.6867	2.7428	0.0109	0.1257	0.0142	0.1393
	4	0.9523	5.0084	0.0064	0.1534	0.0087	0.1613
30	1	1.6643	2.8226	0.0603	0.8850	0.1282	1.7482
	2	0.1416	0.7613	0.0060	0.0848	0.0125	0.1193
	3	0.6515	1.5962	0.0115	0.0705	0.0138	0.0830
	4	0.9504	2.9910	0.0057	0.0910	0.0061	0.1025

以上我们给出了一种定数截尾情形下三参数威布尔分布的矩估计方法. 不完全样本矩估计的大样本性质还有待进一步研究. 这种得出矩方程的思想, 对来自其他分布下的不完全样本的矩估计问题也提供了思路.

感谢审稿人对本文提出的宝贵意见!

参 考 文 献

- [1] Lawless, J.F., *Statistical Models and Methods for Lifetime Data*, John Wiley, 1982.
- [2] 陈迪, 截尾样本的三参数 Weibull 分布参数的近似极大似然估计 (AMLE), *数理统计与应用概率*, **6(3)**(1991), 368-381.
- [3] Cohen, A.C., Whitten, B.J., Modified maximum likelihood and modified moment estimators for the three-parameter Weibull distribution, *Commun. Statist. - Theory Meth.*, **11**(1982), 2631-2656.
- [4] Cohen, A.C., Whitten, B.J. and Ding, Y., Modified Moment Estimation for the Three-parameter Weibull Distribution, *J. Qual. Tech.*, **16**(1984), 159-167.
- [5] 倪中新, 几种不完全数据的统计分析, 上海师范大学硕士论文, 2000.
- [6] Fei, H.L., Su, D.Q. and Leng, S.M., *Statistical Analysis for the Residual Life of Progressive Stress Screening Tests*, Icrecm'92, International Academic Publishers.
- [7] Balakrishnan, N. and Cohen, A.C., *Order Statistic and Inference*, Academic Press Inc., 1991.

Moment Estimation of the Three-parameter Weibull Distribution under Censored Samples

SUN LIBIN

(Xuzhou Normal University, Xuzhou, 221116)

FEI HELIANG

(Shanghai Teachers University, Shanghai, 200234)

Based on censored samples, the problem of moment estimation of the three-parameter Weibull distribution is discussed. mean residual lifetime is used to constructed sample moment. With the method of transformation, the modified moment equation is derived, and the modified moment estimators (MME) are proposed thereby.