

# 三参数 Weibull分布几种常用的参数估计方法

胡恩平, 罗兴柏, 刘国庆

(军械工程学院, 河北 石家庄 050003)

**摘 要:** 介绍了三参数 Weibull分布几种常用的参数估计方法, 并对这些方法的特点做了分析和比较, 以供应用参考。

**关 键 词:** 三参数 Weibull分布; 极大似然估计; Bayes估计; 相关系数优化法; 矩法

中图分类号: TH28

文献标识码: A

三参数 Weibull分布是一种较为完善的分布, 在拟合随机数据时有很大的灵活性和适应性, 当形状参数  $U$  取不同值时, 它可以等效或接近于其它一些常用的分布, 如:  $U=3\sim 4$  时近似于正态分布,  $U=1$  时为指数分布,  $U=2$  时为端利分布. 三参数 Weibull分布也具有局限性, 它对参数的估计比较复杂, 对统计性质的证明比较困难. 近几年来, 研究工作的进一步深入和计算机辅助设计技术的发展, 使得三参数 Weibull分布在工程中的应用日趋广泛.

## 1 极大似然估计法

极大似然估计是一种十分有效和通用的参数估计方法, 在参数估计问题中占有基本的地位, 尤其在处理不完全寿命的情况下, 极大似然估计具有明显的优势. 但是, 在三参数 Weibull分布估计时, 该方法的计算量大, 比较费时, 因而在实际使用中, 人们往往是在极大似然函数的基础上, 结合其它的方法来对三参数 Weibull分布进行分析, 这样既保证了采用极大似然估计的精度, 又避免了繁杂的计算.

设三参数 Weibull分布的分布函数形式为

$$F(x) = 1 - \exp\left[-\left(\frac{x-T}{Z}\right)^U\right] \quad (1)$$

其中,  $T$ ,  $U$ ,  $Z$  分别称为位置参数、形状参数和特征参数. 对于容量为  $n$  的全子样数据  $x_1, x_2, \dots, x_n$ , 参数的似然函数为

$$L(x, T, U, Z) = \prod_{i=1}^n \frac{U}{Z} \left(\frac{x_i - T}{Z}\right)^{U-1} \cdot \exp\left[-\left(\frac{x_i - T}{Z}\right)^U\right] \quad (2)$$

对数似然方程为

$$\frac{\partial \ln L}{\partial U} = \frac{n}{U} + \sum_{i=1}^n \ln \frac{x_i - T}{Z} - \sum_{i=1}^n \left( \frac{x_i - T}{Z} \right)^U \ln \frac{x_i - T}{Z} = 0$$

(3)

$$\frac{\partial \ln L}{\partial Z} = -\frac{n}{Z} - \frac{n(U-1)}{Z} + \frac{U}{Z} \sum_{i=1}^n \left( \frac{x_i - T}{Z} \right)^U = 0$$

(4)

$$\frac{\partial \ln L}{\partial T} = (1-U) \sum_{i=1}^n (x_i - T)^{-1} + \frac{U}{Z} \sum_{i=1}^n \left( \frac{x_i - T}{Z} \right)^{U-1} = 0$$

(5)

这是三个超越方程,需用数值方法迭代求解.由式(4)得

$$Z = \left[ \frac{1}{n} \sum_{i=1}^n (x_i - T)^U \right]^{1/U}$$

(6)

将式(6)代入式(3)和式(5)分别得到

$$\frac{1}{U} - \frac{\sum_{i=1}^n (x_i - T)^U \ln(x_i - T)}{\sum_{i=1}^n (x_i - T)^U} + \frac{1}{n} \sum_{i=1}^n \ln(x_i - T) = 0$$

(7)

$$\frac{U-1}{U} \sum_{i=1}^n (x_i - T)^{-1} + \frac{n \sum_{i=1}^n (x_i - T)^{U-1}}{\sum_{i=1}^n (x_i - T)^U} = 0$$

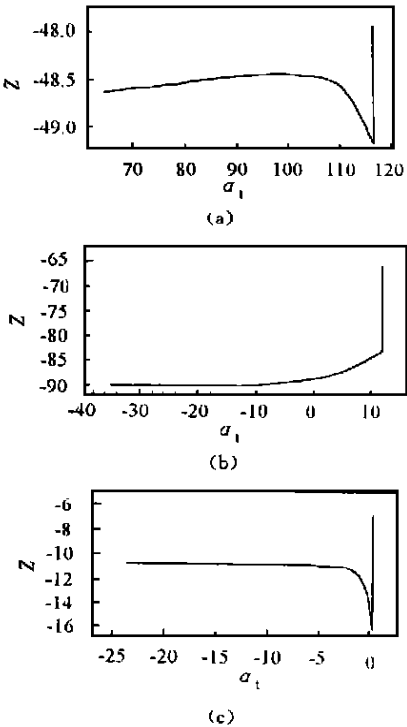
(8)

上述结果是在全子样情况下推出来的,对于截尾样本,用类似的方法也可导出相应的结果.在循环迭代之前先要输入初值,这给无先验信息或先验信息不足的样本数据带来困难,而初值的选择恰当与否将直接影响迭代结果.下面介绍的是借助似然函数曲线图来估计参数的方法.

设  $L^*(T) = L(T, U, Z)$ , 这是关于  $U$  和  $Z$  的极大似然函数,其中  $T$  为相对固定值,  $U, Z$  分别为方程(6)和(7)的解.图 1a 图 1b 和图 1c 分别为对应于表 1 中的第 1 组、第 2 组和第 3 组数据曲线  $Z = \ln L^*(T)$  的三种可能的曲线类型.

表 1 拟合数据

第 1 组	225	171	198	189	189	135	162	135	117	162
第 2 组	74	57	48	29	502	12	70	21	29	386
	153	26	326							
第 3 组	0.273	0.468	0.504	0.535	0.617	0.804	0.932			
	1.034	1.289	1.293	1.294	1.376	1.399	1.407			
	1.422	1.497	1.521	1.542	1.685	1.737				



可以看出,只有图 1a 有局部极大值,而且只有这条曲线通过在  $Z$  出现最大值处  $T$  的值,才能找到真正的极

图 1 对数似然函数曲线图

大似然解.这样,  $U$  的估计值可以从式 (7) 和式 (8) 得到,  $Z$  的估计值可以用  $T$  和  $U$  的估计值从式 (6) 中求得.在图 1a 中的曲线上,  $Z$  的最小值为似然函数提供驻点. Rockette<sup>[1]</sup> 等人论证过, 当  $U > 1, n \rightarrow \infty$  时, 图 1a 出现的概率为 1, 他们解释说局部最大值的出现意味着驻点的存在. 图 1b 中,  $Z$  随  $T$  的减小而下降, 对无穷样本 (即  $n \rightarrow \infty$ ), 当  $U < 1$  时, 该曲线出现的概率为 1, 则  $T$  的极大似然估计值  $T_{MLE} = x(1)$  ( $x(1)$  为顺序统计量的最小值). 图 1c 中  $Z$  出现极小值, 但没有极大值, 因而无极大似然解.

图 1 中的极大值极小值对应着方程 (7)、(8) 中  $U$  和  $T$  的解以及方程 (6) 中  $Z$  的解, 当  $T$  取遍  $T$  值, 由方程 (7)、(8) 可以得到两条曲线  $P_1(U \sim T)$  和  $P_2(U \sim T)$ . 根据  $P_1$  和  $P_2$  交点的不同情况, 相应地就可以确定解的情况<sup>[2]</sup>: ① 无交点, 则无极大似然解; ② 一个交点, 无极大似然解, 应改用极值分布来拟合; ③ 两个交点, 对应于较小的  $T$  为极大似然解.

## 2 Bayes 统计分析

虽然目前统计学界对 Bayes 统计还存在分歧, 但人们依然利用 Bayes 统计方法解决了很多实际问题. Bayes 统计推断的原理可以这样来理解: 一经由样本算出了未知参数的后验分布, 其余的东西 (样本值、样本分布、先验分布) 就全忘记了, 这时对未知参数作判断的唯一凭借就是后验分布. Bayes 学派虽然也用到样本分布, 但只是利用分布的数学形式, 将其与先验分布配合而达到后验分布.

将式 (1) 写成如下的形式

$$F(x, T, U, \theta) = - \exp \left[ - \frac{1}{\theta} (x - T) \right] \quad (9)$$

设满足分布函数式 (9) 的受试样本  $T_1, T_2, \dots, T_n$  独立同分布, 其截尾时间  $\{L_i\}$  相互独立, 分布函数为  $G(x)$ , 密度函数为  $g(x)$ ,  $G(x)$ 、 $g(x)$  均与参数  $T, U, \theta$  无关, 且  $\{T_i\}$  与  $\{L_i\}$  也相互独立. 假定只能观测到  $X_i = \min\{T_i, L_i\}, i = 1, 2, \dots, n$ , 记  $W = 1$  (当  $T_i \leq L_i$  时),  $W = 0$  (当  $T_i > L_i$  时), 则  $(X_i, W)$  相互独立, 故似然函数为

$$L(T, U, \theta) = \prod_{i=1}^n [g(x_i)]^{1-W_i} [1 - G(x_i)]^{W_i} \prod_{i=1}^n [f(x_i)]^{W_i} [1 - F(x_i)]^{1-W_i}$$

其中  $f(x)$  为  $F(x)$  的密度函数,  $x_i$  为  $X_i$  的观察值. 取  $(T, U, \theta)$  的先验分布<sup>[3]</sup> 为  $\Delta(T, U, \theta) \propto 1/(\theta U)$ , 从而可知  $(T, U, \theta)$  的后验分布为

$$H(T, U, \theta) \propto \left[ \frac{U}{\theta} \right]^{W-1} \theta^{-2} \left[ \prod_{i=1}^n (x_i - T)^{W_i} \right]^{U-1} \exp \left[ - \frac{1}{\theta} \sum_{i=1}^n (x_i - T) \right] \quad (10)$$

式中

$$W = \sum_{i=1}^n W_i$$

给定  $x, W$  时,  $U, \theta, T$  的后验边缘分布分别为

$$H_1(U) = A_1 \Gamma(W) \int_0^{T(1)} U^W \left[ \prod_{i=1}^n (x_i - T) \right]^{-W} \left[ \prod_{i=1}^n (x_i - T)^{W_i} \right]^{U-1} dT = A_1 \Gamma(W) B_1(U) \quad (11)$$

$$H_2(\theta) = A_2 \theta^{-(W-1)} \int_0^{T(1)} \int_0^{+\infty} U^W \left[ \prod_{i=1}^n (x_i - T)^{W_i} \right]^{U-1} \exp \left[ - \frac{1}{\theta} \sum_{i=1}^n (x_i - T) \right] dU \cdot dT = A_2 B_2(\theta)$$

$$H_3(T) = A_3 \Gamma(W) \int_0^{+\infty} U^{W-1} \left[ \sum_{i=1}^n (x_i - T) \right]^{-W} \left[ \prod_{i=1}^n (x_i - T)^{W_i} \right]^{U-1} dU = A_3 \Gamma(W) B_3(T) \quad (13)$$

其中

$$\begin{aligned} T(1) &= \min\{x_i\} \\ A_1^{-1} &= \Gamma(W) \int_0^{+\infty} B_1(U) dU \\ A_2^{-1} &= \int_0^{+\infty} B_2(\theta) d\theta \\ A_3^{-1} &= \Gamma(W) \int_0^{T(1)} B_3(T) dT \end{aligned}$$

在平方损失下,  $U, \theta, T$  的 Bayes 的点估计分别为

$$\begin{aligned} \hat{U} &= E(U) = A_1 \Gamma(W) \int_0^{+\infty} U B_1(U) dU \\ \hat{\theta} &= E(\theta) = A_2 \int_0^{+\infty} \theta B_2(\theta) d\theta \\ \hat{T} &= E(T) = A_3 \Gamma(W) \int_0^{T(1)} T B_3(T) dT \end{aligned}$$

$U, \theta, T$  的  $(1 - V)$  Bayes 置信区间  $(U_L, U_H)$ ,  $(\theta_L, \theta_H)$ ,  $(T_L, T_H)$  分别由以下各式决定

$$A_1 \Gamma(W) \int_0^{U_L} B_1(U) dU = A_1 \Gamma(W) \int_{U_H}^{+\infty} B_1(U) dU = V/2 \quad (14)$$

$$A_2 \int_0^{\theta_L} B_2(\theta) d\theta = A_2 \int_{\theta_H}^{+\infty} B_2(\theta) d\theta = V/2 \quad (15)$$

$$A_3 \Gamma(W) \int_0^{T_L} B_3(T) dT = A_3 \Gamma(W) \int_{T_H}^{T(1)} B_3(T) dT = V/2 \quad (16)$$

### 3 相关系数优化法

对可靠度函数

$$R(x) = \exp \left[ - \left( \frac{x - T}{Z} \right)^U \right] \quad (17)$$

作如下变换:

$$Y = \ln \left[ - \ln R(x) \right], X = \ln(x - T), A = -U \ln Z, B = U \quad (18)$$

可得

$$Y = A + BX \quad (19)$$

式 (19) 中变量  $X$  与  $Y$  之间成线性关系, 根据观测数据  $(x_i, R(x_i))$ , 通过式 (18) 得到新的数据  $(X_i, Y_i)$ , 于是由线性回归分析得到待估参数  $A, B$  和线性相关系数  $r$  分别为

$$\begin{aligned} \hat{A} &= \bar{Y} - \hat{B} \bar{X} \\ \hat{B} &= \frac{L_{XY}}{L_{XX}} \\ r &= \frac{L_{XY}}{\sqrt{L_{XX} L_{YY}}} \end{aligned} \quad (20)$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$$

$$L_{XX} = \sum_{i=1}^n X_i^2 - n\bar{X}^2$$

$$L_{YY} = \sum_{i=1}^n Y_i^2 - n\bar{Y}^2$$

$$L_{XY} = \sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y}$$

由上式可见  $A$ 、 $B$  和  $r$  为  $T$  的函数, 记为  $A(T)$ 、 $B(T)$  和  $r(T)$ . 我们所要求的  $T$  必须使线性相关系数  $r$  的绝对值  $|r|$  取最大值, 由  $d|r(T)|/dT = 0$  可求得  $T$ , 等价于  $dr^2(T)/dT = 0$ . 因为

$$\frac{dr^2}{dT} = 2r^2 \left[ \frac{1}{L_{XY}} \frac{dL_{XY}}{dT} - \frac{1}{2L_{XX}} \frac{dL_{XX}}{dT} \right] \quad (21)$$

$$\text{故可得} \quad (1/L_{XY})(dL_{XY}/dT) - [1/(2L_{XX})] (dL_{XX}/dT) = 0 \quad (22)$$

上式不便直接求解, 必须通过数值计算方法求解. 令

$$L_{X0} = - \frac{dL_{XX}}{2dT} = - \frac{1}{n} \sum_{i=1}^n X_i \sum_{i=1}^n (x_i - T)^{-1} + \sum_{i=1}^n X_i (x_i - T)^{-1}$$

$$L_{Y0} = - \frac{dL_{XY}}{dT} = - \frac{1}{n} \sum_{i=1}^n Y_i \sum_{i=1}^n (x_i - T)^{-1} + \sum_{i=1}^n Y_i (x_i - T)^{-1}$$

代入式 (22) 得

$$(L_{X0}/L_{XX}) - (L_{Y0}/L_{XY}) = 0 \quad (23)$$

令  $H(T) = (L_{X0}/L_{XX}) - (L_{Y0}/L_{XY})$ , 设  $T_e$  为  $T$  的预估值, 由以上推导可知,  $H(T)$  与  $dr^2/dT$  同号, 且当  $T_e < T$  时,  $(dr^2/dT)|_{T_e} > 0$ , 即  $H(T_e) > 0$ ; 当  $T_e > T$  时,  $(dr^2/dT)|_{T_e} < 0$ , 即  $H(T_e) < 0$ . 于是可以按如下的方法来计算: 先求  $H(0)$ , 因为  $T$  位于区间  $[0, x_{(1)}]$  内,  $x_{(1)}$  为顺序统计量的最小值, 如果  $H(0) \leq 0$ , 取  $T = 0$ ; 如果  $H(0) > 0$ , 则按半分法进行编程计算, 直到满足所要求的精度求得  $T$ . 求出  $T$  后, 由式 (20) 求得  $A$ 、 $B$  和  $r$ , 再由式 (18) 求得  $U$  和  $Z$  的值.

## 4 矩法

设  $X$  服从式 (1) 分布, 则数学期望、方差和偏度分别为

$$E(x) = T + Zg_1(U) \quad (24)$$

$$D(x) = Z[g_2(U) - g_1^2(U)] \quad (25)$$

$$B(x) = \frac{g_3(U) - 3g_2(U)g_1(U) + 2g_1^3(U)}{[g_2(U) - g_1^2(U)]^{3/2}} \quad (26)$$

其中  $g_i(U) = \Gamma(1 + i/U)$ . 设试验数据为  $x_1, x_2, \dots, x_n$ , 则可求得样本均值、方差和偏度分别为

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad B_0 = \frac{n \sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)(n-2)S^3} \quad (27)$$

将由式 (27) 求出的样本均值、方差和偏度, 分别作为  $E(x)$ 、 $D(x)$  和  $B(x)$  的点估计, 然后由半分法, 从式 (26) 迭代出  $U$ , 再由式 (24)、(25) 求出  $Z$  和  $T$ .

如果先求出位置参数  $T$ , 由式 (24)、(25) 也能方便地求出  $U$  和  $Z$ . 利用三点作二次插值函数

$$P(x) = kx^2 + bx + c \quad (28)$$

其中

$$k = \frac{(x_2 - x_3)F(x_1) + (x_3 - x_1)F(x_2) + (x_1 - x_2)F(x_3)}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)}$$

$$b = \frac{(x_2^2 - x_3^2)F(x_1) + (x_3^2 - x_1^2)F(x_2) + (x_1^2 - x_2^2)F(x_3)}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)}$$

$$c = \frac{(x_3 - x_2)x_2x_3F(x_1) + (x_1 - x_3)x_1x_3F(x_2) + (x_2 - x_1)x_1x_2F(x_3)}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)}$$

式 (28) 对应方程  $kx^2 + bx + c = 0$  小于  $x_1$  的根即为  $T$  的值。

## 5 结束语

本文介绍了四种有关三参数 Weibull 分布的分析方法, 可为研究工作提供参考。一般情况下, 极大似然法的精度较高, 作图法和矩法的精度低一些, 但在实际应用中选择哪种方法, 应根据精度、计算量等要求来决定。在没有具体限制条件的情况下, 可以通过比较来选择合适的分析方法。

## 参考文献:

- [1] Rockette H, Antle C, Klimko L A. Maximum likelihood estimation with the weibull model[J]. J Am Statist Assoc, 1974, 69(3).
- [2] Richard A Lockhart, Michael A Stephens. Estimation and test-of-fit for the three parameter weibull distribution[R]. Technical Report No. 476, Stanford, California, 1993.
- [3] Sinha S K, Slon J A. Bayes estimators of the parameters of the parameters and reliability function of the 3-parameter weibull distribution[J]. IEEE transactions on Reliability, 1988, 37(4).
- [4] 谷耀新. 三参数威布尔分布参数估计方法 [J]. 沈阳工业学院学报, 1997, 16(4).
- [5] 师义民. 定数截尾寿命试验三参数威布尔分布的 Bayes 统计分析 [J]. 工程数学学报, 1992, 19(3).
- [6] 吴云, 宋乾坤. 三参数 Weibull 分布下随机截尾恒加寿命试验的 Bayes 统计分析 [J]. 西南民族学院学报 (自然科学版), 1997, 23(2).
- [7] 蒋仁言. 威布尔模型族特性: 参数设计及应用 [M]. 北京: 科学出版社, 1998.
- [8] Adatia A, Chan L K. Robust estimators of the 3-parameter weibull distribution[J]. IEEE transactions on Reliability, 1985, 34(3).
- [9] Newby, Martin. Properties of moment estimators for the 3-parameter weibull distribution [J]. IEEE transactions on Reliability, 1984, 33(2).
- [10] Garn G W. Moment estimators for the 3-parameter weibull distribution [J]. IEEE transactions on Reliability, 1988, 37(3).
- [11] A D Graham. Experinec with two estimators for the three-parameter weibull distribution structures note 484 [Z]. Victoria, Australia, 1983.
- [12] 曲延禄, 张程道, 阎书源. 三参数 weibull 分布的参数估计 [J]. 气象学报, 1987(3).

## Parameter Estimating Methods for the Three Parameters Weibull Distribution

HU En-ping, LUO Xing-bai, LIU Guo-qing

(Ordnance Engineering College, Shijiazhuang 050003, China)

**Abstract** This paper introduces several parameter estimating methods in common use for the three parameters Weibull distribution. Analysis and comparsion to the characteristics of these methods is made so as to reference in application.

**Key words** three parameters Weibull distribution; maxmium likelihood estimation; Bayes estimation; relative coefficient optimization method; moment method