极大似然估计与最优化数值方法

陈怡南

(连云港化工高等专科学校,连云港,222001)

摘 要 本文研究了可靠性分析中基于不完全样本数据寿命分布参数的极大似然估计(MLE)的最优化数值解法,并且对威布尔(Weiull)分布和对数正态分布分别给出了具体的随机模拟计算结果。

关键词 极大似然估计 最优化 随机模拟

1 问题的提出

在可靠性分析中,寿命数据分析是一项基础性工作。用 X 表示某种产品的寿命,其分布函数是 $F(x;\theta)$,$\theta\in\Theta$,这里 θ 是未知参数,Θ 是非空开集。从这种产品中随机抽取 n 件进行寿命试验,如果试验一直进行到所有 n 件产品都失效为止,那么便得到一组完全寿命样本数据 x_1, $x_2,\cdots,x_n$。然而,在工程实际和许多研究中,由于种种条件的限制和诸多因素的干扰,有时不可能获得完全样本。如,受试验时间、费用等限制,不可能将寿命试验做到所有试件都失效。又如,受试验环境、仪器设备条件的影响,以及观测过程中的随机干扰,也不可能得到完整的寿命数据。这时我们只能得到一组不完全寿命样本数据。就试验类型而言,不完全寿命样本主要产生于以下几种截尾寿命试验模型:定数截尾试验,定时截尾试验和随机截尾试验等。

本文考虑如下带有不完全信息的随机截尾试验模型[1]。

设 $x_1,x_2\cdots$ 是来自总体 $F(x;\theta)$ 的寿命样本,每个 x_i 的密度函数为 $f(x;\theta)$。又设截尾时间 $y_1,y_2\cdots$ 是相互独立取正值的随机变量序列,y_i 的分布函数是 $G_i(y)$,$i=1,2\cdots$,它们与参数 θ 无关。假定 $\{x_i\}$ 与 $\{y_i\}$ 相互独立。对 n 个受试样本,令

$$\alpha_i=\begin{cases}1,\text{若 } x_i<y_i\\0,\text{若 } x_i\geqslant y_i\end{cases},\qquad \beta_i=\begin{cases}0,\text{若 } x_i<y_i \text{ 且失效未被显示}\\1,\text{其它}\end{cases}$$

并设

$$P\{\beta_i=1|\alpha_i=1\}=p,\qquad P\{\beta_i=0|\alpha_i=1\}=1-p,\quad (0<p\leqslant 1)$$

其中 p 表示失效显示概率。又令:当 $\alpha_i=1,\beta_i=1$ 时,$z_i=x_i$;当 $\alpha_i=1,\beta_i=0$ 时,$z_i=y_i$;当 $\alpha_i=0$(隐含 $\beta_i=1$)时,$z_i=y_i$;$i=1,2,\cdots,n$,则 $\{(z_i,\alpha_i,\beta_i),1\leqslant i\leqslant n\}$ 就是我们获得的带有不完全信息的随机截尾数据。基于数据 $\{(z_i,\alpha_i,\beta_i)\}$ 的似然函数为

$$L(\theta)=A\prod_{i=1}^{n}[f(z_i;\theta)]^{\alpha_i\beta_i}[F(z_i;\theta)]^{\alpha_i(1-\beta_i)}[\bar{F}(z_i;\theta)]^{1-\alpha_i} \tag{1}$$

其中,A 是与参数 θ 无关的正函数。

相应的对数似然函数为

$$\ln L(\theta)=\sum_{i=1}^{n}[\alpha_i\beta_i\ln f+\alpha_i(1-\beta_i)\ln F+(1-\alpha_i)\ln\bar{F}] \tag{2}$$

这里略去了与 θ 无关的 $\ln A$ 这一项。

为了得到参数 θ 的极大似然估计(MLE),需要求出 $\ln L(\theta)$ 的极大值点。经过适当的转换,可以把求 MLE 的问题转化为某一形式的最优化求解问题。

本文的以下部分将针对威布尔分布和对数正态分布这两种常见的寿命分布,通过随机模拟产生不完全寿命样本数据$\{(z_i,\alpha_i,\beta_i),1\leqslant i\leqslant n\}$,并运用无约束最优化数值方法求出分布参数的 MLE。

2 威布尔分布参数的 MLE

设产品寿命 X 服从威布尔分布,其分布函数与密度函数分别为

$$F(x;\lambda,b)=1-\exp(-\lambda x^b),\quad f(x;\lambda,b)=\lambda b x^{b-1}\exp(-\lambda x^b)\ (x>0)$$

其中 λ,b 是未知的正参数。

由(2)可得对数似然函数 $\ln L(\lambda,b)$,将它分别对 λ、b 求偏导数并令其为零,再经过一定的整理,得关于参数 λ、b 的似然方程组为

$$\begin{cases}\sum_{i=1}^{n}\left[\alpha_i\beta_i+\alpha_i(1-\beta_i)\dfrac{\lambda z_i^b}{\exp(\lambda z_i^b)-1}-(1-\alpha_i+\alpha_i\beta_i)\lambda z_i^b\right]=0\\ \sum_{i=1}^{n}\left[\alpha_i\beta_i+\alpha_i(1-\beta_i)\dfrac{\lambda z_i^b}{\exp(\lambda z_i^b)-1}-(1-\alpha_i+\alpha_i\beta_i)\lambda z_i^b\right]\ln z_i+\dfrac{1}{b}\sum_{i=1}^{n}\alpha_i\beta_i=0\end{cases}\tag{3}$$

可以证明[2],该方程组在一定条件下有唯一的一组解 $\hat{\lambda}_n$、$\hat{b}_n$,并且它们分别为 λ、b 的 MLE。因此,下面通过随机模拟的方法产生数据$(z_i,\alpha_i,\beta_i)(i=1,2,\cdots,n)$,并求解似然方程组(3),从而得到 $\hat{\lambda}_n$ 和 $\hat{b}_n$。

为了求解非线性方程组(3),下面采用无约束最优化梯度法。取目标函数为

$$H(\lambda,b)=h_1^2(\lambda,b)+h_2^2(\lambda,b)$$

其中

$$h_1=\sum_{i=1}^{n}\left[\alpha_i\beta_i+\alpha_i(1-\beta_i)\frac{\lambda z_i^b}{\exp(\lambda z_i^b)-1}-(1-\alpha_i+\alpha_i\beta_i)\lambda z_i^b\right]$$

$$h_2=\sum_{i=1}^{n}\left[\alpha_i\beta_i+\alpha_i(1-\beta_i)\frac{\lambda z_i^b}{\exp(\lambda z_i^b)-1}-(1-\alpha_i+\alpha_i\beta_i)\lambda z_i^b\right]\ln z_i+\frac{1}{b}\sum_{i=1}^{n}\alpha_i\beta_i$$

则求解(3)就可以转化为求解无约束最优化问题

$$\min H(\lambda,b)$$

由于 $H(\lambda,b)\geqslant 0$,而方程组(3)有唯一的一组解,因此,存在唯一的点$(\hat{\lambda}_n,\hat{b}_n)$,使 $H(\lambda,b)$在该点达到最小值 0。

现选定一组(λ,b)的初值,计算对应的目标函数值 H。若 $H<\varepsilon$,即 H 满足精度要求,则该组初值即为所求的解。否则,从该点出发,在负梯度方向上确定一组值替代初值,以使函数值 H 下降得最快。重复上述步骤,直到 H 满足精度要求为止。

我们对(λ,b)的三组不同取值情况进行模拟计算作为例子,下表列出了其中一组的计算结果。在这里,样本容量分别取 $n=20,30,50,60,80$;失效显示概率分别取 $p=0.2,0.5,0.8,1.0$;截尾模型分别选取四种:指数分布 $E_x(0.5)$、$E_x(1.0)$、$E_x(1.5)$和均匀分布 $U(0.5,1.5)$。模拟计算遍数取 50,表中列出的结果是 50 遍求解结果的平均值。

表　威布尔分布参数 λ 和 b 的 MLE

$\lambda=1.0, b=0.5$

n	Censoring Pattern	$p=0.2$		$p=0.5$		$p=0.8$		$p=1.0$	
		$\hat{\lambda}_n$	$\hat{b}_n$	$\hat{\lambda}_n$	$\hat{b}_n$	$\hat{\lambda}_n$	$\hat{b}_n$	$\hat{\lambda}_n$	$\hat{b}_n$
20	$Ex(0.5)$	1.0074	0.5977	0.9590	0.5945	0.9635	0.5843	0.9529	0.5772
	$Ex(1.0)$	1.0607	0.6472	1.0080	0.6539	1.0053	0.5947	0.9551	0.5870
	$E(x)(1.5)$	1.0359	0.6238	1.0021	0.5998	0.9561	0.5972	0.9307	0.5896
	$U(0.5,1.5)$	1.0939	0.7461	1.0595	0.6226	1.0139	0.5978	0.9755	0.5736
30	$Ex(0.5)$	1.0130	0.5820	1.0275	0.5584	1.0169	0.5456	1.0021	0.5333
	$Ex(1.0)$	1.1020	0.5694	1.0769	0.5641	1.0437	0.5427	0.983	0.5410
	$Ex(1.5)$	1.1111	0.6041	1.0613	0.5852	0.9838	0.5676	0.9623	0.5607
	$U(0.5,1.5)$	1.0211	0.6421	1.0172	0.5569	1.0052	0.5568	1.0006	0.5280
50	$Ex(0.5)$	1.0138	0.5317	1.0015	0.5215	0.9894	0.5286	0.9902	0.5206
	$Ex(1.0)$	1.0273	0.5253	1.0165	0.5232	1.0029	0.5110	0.9899	0.5238
	$Ex(1.5)$	1.0510	0.5219	1.0096	0.5237	1.0098	0.5333	0.9787	0.5320
	$U(0.5,1.5)$	1.0267	0.5483	1.0254	0.5296	1.0057	0.5170	0.9929	0.5142
60	$Ex(0.5)$	1.0027	0.5814	1.0121	0.5461	1.0092	0.5355	0.9976	0.5289
	$Ex(1.0)$	1.0185	0.5358	1.0153	0.5312	0.9975	0.5270	0.9857	0.5270
	$Ex(1.5)$	1.0236	0.5314	1.0129	0.5226	0.9905	0.5236	0.9716	0.5335
	$U(0.5,1.5)$	0.9870	0.5914	0.9896	0.5237	0.9909	0.5273	0.9872	0.5162
80	$Ex(0.5)$	1.0005	0.5230	0.9870	0.5253	0.9882	0.5277	0.9871	0.5226
	$Ex(1.0)$	0.9960	0.5237	0.9946	0.5094	0 9858	0.5198	0.9860	0.5246
	$Ex(1.5)$	0.9903	0.5356	0.9962	0.5353	0.9647	0.5281	0.9434	0.5202
	$U(0.5,1.5)$	1.0086	0.5259	1.0083	0.5372	1.0091	0.5365	0.9923	0.5285

3　对数正态分布参数的 MLE

设 X 服从对数正态分布，其密度函数为

$$f(x;\mu,\sigma)=\frac{1}{\sqrt{2\pi}\sigma x}\exp\left[-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right]\quad(x>0)$$

分布函数为

$$F(x;\mu,\sigma)=\int_0^x f(t;\mu,\sigma)\mathrm{d}t=\Phi\left(\frac{\ln x-\mu}{\sigma}\right)\quad(x>0)$$

其中，μ、σ 是未知参数且 $-\infty<\mu<+\infty$，$\sigma>0$，Φ 是标准正态分布 $N(0,1)$ 的分布函数。

由(2)可得基于数据 $\{(z_i,\alpha_i,\beta_i),\ 1\leqslant i\leqslant n\}$ 的关于参数 μ 和 σ 的对数似然函数为

$$\ln L(\mu,\sigma)=\sum_{i=1}^{n}[\alpha_i\beta_i\ln f(z_i;\mu,\sigma)+\alpha_i(1-\beta_i)\ln F(z_i;\mu,\sigma)+(1-\alpha_i)\ln\bar{F}(z_i;\mu,\sigma)]\quad(4)$$

欲求 $\ln L(\mu,\sigma)$ 的极大值点，也就是要求负对数似然函数 $B(\mu,\sigma)=-\ln L(\mu,\sigma)$ 的极小值点，即求解下列最优化问题：

$$\min B(\mu,\sigma)\quad(5)$$

由(4)式及对数函数的性质，得

$$B(\mu,\sigma)=\sum_{i=1}^{n}[\alpha_i\beta_i\ln f^{-1}(z_i;\mu,\sigma)-\alpha_i(1-\beta_i)\ln F(z_i;\mu,\sigma)-(1-\alpha_i)\ln\bar{F}(z_i;\mu,\sigma)]$$

由于 $f^{-1}=\sqrt{2\pi}\sigma z_i\cdot\exp\left[\frac{1}{2}\left(\frac{\ln z_i-\mu}{\sigma}\right)^2\right]$，故

$$B(\mu,\sigma)=\sum_{i=1}^{n}\left[\alpha_i\beta_i\left(\ln(\sqrt{2\pi}z_i)+\ln\sigma+\frac{1}{2}\left(\frac{\ln z_i-\mu}{\sigma}\right)^2\right)\right.$$

$$\left.-\alpha_i(1-\beta_i)\ln\Phi\left(\frac{\ln z_i-\mu}{\sigma}\right)-(1-\alpha_i)\ln\left(1-\Phi\left(\frac{\ln z_i-\mu}{\sigma}\right)\right)\right] \tag{6}$$

下面通过随机模拟的方法来产生数据$\{(z_i,\alpha_i,\beta_i),1\leqslant i\leqslant n\}$，并运用无约束最优化直接法来求解问题(5)。

首先，给定一组参数值(μ,σ)和样本容量值n(本文分别取$n=30,40,50,60,70,80$)，并且取截尾模型是均匀分布$\mu(0.5,4.5)$。

其次，根据给定的μ和σ，用蒙特卡罗(Monte－Carlo)方法产生对数正态分布随机数$x_1,x_2,\cdots,x_n$(寿命样本值)，再产生区间(0.5,4.5)上的截尾数据$y_1,y_2,\cdots,y_n$；根据已产生的$\{x_i\}$、$\{y_i\}$和取定的失效显示概率值p(本节分别取$p=0.5,0.8,1.0$)，经过逐对比较x_i和y_i的大小，按以下做法便确定出带有不完全信息的随机截尾数据$\{(z_i,\alpha_i,\beta_i),1\leqslant i\leqslant n\}$：

当$x_i\geqslant y_i$时，得$z_i=y_i,\alpha_i=0,\beta_i=1$；

当$x_i<y_i$时，产生一个(0,1)上的均匀分布随机数r，若$r\leqslant p$，则得$z_i=x_i,\alpha_i=1,\beta_i=1$；若$r>p$则得$z_i=y_i,\alpha_i=1;\beta_i=0$。

最后，将数据$\{(z_i,\alpha_i,\beta_i),1\leqslant i\leqslant n\}$代入(6)式，求解最优化问题(5)。

为了避免较复杂的求导运算，减少计算量，选用单纯形下降法来直接求解最优化问题(5)。

选定一组(μ,σ)的初值，确定初始单纯形(对应于二维空间，单纯形是三角形)，计算各顶点处的目标函数B的值，比较大小，丢掉函数值最大的点，代之以新的点，构成新的单纯形，反复迭代使顶点处的函数值逐步下降，顶点逐步逼近函数的最小值点。

我们对$\mu=0,\sigma=1.0$和$\mu=2.0,\sigma=1.0$这两种情况进行模拟计算。结果表明，MLE与真值的偏差比较小，且MLE的值比较稳定，其精度比较高。

参考文献

1 Elperin, T. and Gertsbakh, I., Estimation in a rendom censoring model with incomplet information: exponential lifetime distribution, IEEE Trans. Rel. 1988, 37(2)

2 陈怡南．带有不完全信息随机截尾试验模型与无失效数据统计分析研究．南京航空航天大学硕士论文．1995

3 王子若等．优化计算方法．机械工业出版社．1989