文章编号: 1001-4543(2021)01-0056-05 DOI: 10.19570/j.cnki.jsspu.2021.01.008

逐步增加定数截尾下 Gompertz 分布的统计分析

李 琼 [1], 武 东 [2]

(1. 徽商职业学院 电子信息系, 合肥 231201; 2. 安徽农业大学 理学院, 合肥 230036)

摘 要: 讨论了基于逐步增加定数截尾样本 Gompertz 分布的参数估计, 提出了模型参数的逆矩估计和一种新的联合置信区间。最后通过 Monte Carlo 方法进行了模拟研究, 统计表明方法均有效。

关键词: Gompertz 分布; 逐步增加定数截尾; 逆矩估计; 区间估计

中图分类号: F222.1 **文献标志码**: A

Statistical Analysis of Gompertz Distribution under Progressive Type-II Censoring

LI Qiong[1], WU Dong[2]

(1. Department of Electronic Information, Huishang Vocational College, Hefei 231201, China;
2. School of Science, Anhui Agricultural University, Hefei 230036, China)

Abstract: The parameter estimation based on the Gompertz distribution of progressive Type-II censorings is discussed, and the inverse moment estimator and a new joint confidence interval estimate are proposed. Finally, it is seen that the estimates are efficient using Monte Carlo method to simulate the parameters of the model.

Keywords: Gompertz distribution; progressive Type-II censoring; inverse moment estimate; interval estimate

0 引言

假设产品的寿命服从两参数 Gompertz 分布, 其分布函数和密度函数分别为:

$$F(x)=1-\exp\left\{-\frac{\theta}{\beta}(\mathrm{e}^{\beta x}-1)\right\},\ x>0 \quad (1)$$

$$f(x)=\theta\mathrm{e}^{\beta x}\exp\left\{-\frac{\theta}{\beta}(\mathrm{e}^{\beta x}-1)\right\},\ x>0 \quad (2)$$

式中: x 通常为寿命, 不能为负值; $\theta>0$ 为尺度参数; $\beta\in R$ 为形状参数。当 $\beta>0$ 时, 失效率函数 $\lambda(x)=\theta\mathrm{e}^{\beta x}$ 为单调递增; 当 $\beta<0$ 时, 失效率函数 $\lambda(x)$ 单调递减。

Gompertz 分布在可靠性与生存分析等领域有着广泛的应用, 周会会 [1] 研究了 Gompertz 分布尺度参数的最优置信区间。Wu 等 [2] 研究了首次失效样本情形下 Gompertz 分布的参数估计。Wu 等 [3] 研究了基于逐步增加定数截尾下 Gompertz 分布的最大似然估计, 并运用 F 分布构造了两参数的联合置信区间。Lenart[4] 研究了 Gompertz 分布的矩以及参数的最大似然估计。Sanku 等 [5] 研究了 Gompertz 分布的各种估计方法和统计性质。而关于 Gompertz 分布的推广主要有: 三参数指数型 Gompertz 分布、逆 Gompertz 分布、Alpha 幂 Gompertz 分布、Unit-Gompertz 分布等 [6-9]。关于 Gompertz 分布在可靠性统计方面的研究, Ismial[10] 研究了在定时截尾

收稿日期: 2020-10-15
通信作者: 武 东 (1976–), 男, 安徽六安人, 副教授, 硕士, 主要研究方向为应用统计、金融时间序列分析。
E-mail: wudong@ahau.edu.cn
基金项目: 安徽省级质量工程项目 (2020jyxm1803), 安徽省高校优秀青年人才支持计划 (gxyq2019254) 资助

下 Gompertz 分布场合部分加速寿命试验的最优设计。Saracoğlu 等 [11] 研究了 Gompertz 分布产品应力强度模型的最大似然估计和区间估计。王蓉华等 [12] 研究了 TFR 模型 Gompertz 分布场合多步步进应力加速寿命试验的最大似然估计。Essam[13] 研究了逐步增加定数截尾情况下广义 Gompertz 分布的某些寿命参数的估计。截至目前, 尚未有文献研究基于逐步增加定数截尾样本 Gompertz 分布参数的逆矩估计, 鉴于此, 本文研究了逐步增加定数截尾样本下 Gompertz 分布场合的逆矩估计, 并提出一种参数联合置信区间的新方法, 最后通过 Monte Carlo 模拟和实例说明得出算法的有效性。

1　参数估计

假设有 n 个产品进行寿命试验, 在第 1 个失效时刻 X_1, 从尚未失效的 $n-1$ 产品中随机选择 R_1 个产品移离试验, 在第 2 个失效时刻 X_2, 从尚未失效的 $n-2-R_1$ 个产品中随机选择 R_2 个产品移离试验, 如此直到第 m 个失效时刻 X_m, 所有尚未失效的 $R_m = n-m-\sum_{i=1}^{m-1} R_i$ 个产品均移离试验, 称为逐步增加定数截尾试验, 得到的随机样本称为逐步增加定数截尾样本, 当 $R_1 = R_2 = \cdots = R_{m-1} = 0$, $R_m = n-m$ 时退化为定数截尾寿命试验。

对于预先给定的移离部件数 ($R_1 = r_1, \cdots, R_{m-1} = r_{m-1}$), 令 $X_1 < X_2 < \cdots < X_m$ 表示来自两参数 Gompertz 分布的逐步增加定数截尾样本。作以下变量变换:

$$Y_i = \frac{\theta}{\beta}(\mathrm{e}^{\beta X_i} - 1), \quad i = 1, 2, \cdots, m \tag{3}$$

式中: $Y_1 < Y_2 < \cdots < Y_\mathrm{m}$ 为来自标准指数分布的逐步增加定数截尾样本; i 为失效时刻的次数。考虑下列变换:

$$\left.\begin{aligned} &Z_1 = nY_1, \\ &Z_2 = (n-r_1-1)(Y_2-Y_1), \\ &\quad\vdots \\ &Z_m = \left(n - \sum_{i=1}^{m-1}(r_i+1)\right)(Y_m - Y_{m-1}) \end{aligned}\right\} \tag{4}$$

Thomas 等 [14] 证明了 $Z_1, Z_2, \cdots, Z_m$ 为独立同分布于标准指数分布。r_i 为第 i 次移离的部件数。

引理 1[15]　设 $Z_1, Z_2, \cdots, Z_m$ 为来自标准指数分布的简单随机样本, 令

$$T_i = \sum_{j=1}^{i} Z_j, \quad i = 1, 2, \cdots, m \tag{5}$$

式中: $\dfrac{T_1}{T_m} \leqslant \dfrac{T_2}{T_m} \leqslant \cdots \leqslant \dfrac{T_{m-1}}{T_m}$ 是来自均匀分布 $U(0,1)$ 的样本容量为 $m-1$ 的次序统计量。

引理 2　假设同上, 记

$$h(\beta) = 2\sum_{i=1}^{m-1} \ln\left(\frac{T_m}{T_i}\right) \tag{6}$$

则:

(1) $h(\beta)$ 服从自由度为 $2m-2$ 的 χ^2 分布。

(2) $h(\beta)$ 在区间 $(-\infty, \infty)$ 是关于 β 的严格增函数。

(3) $\lim\limits_{\beta\to-\infty} h(\beta) = -\infty$, $\lim\limits_{\beta\to+\infty} h(\beta) = +\infty$, 从而对任意 $t \in R$, 方程 $h(\beta) = t$ 在 $(-\infty, \infty)$ 有唯一根。

证明:

(1) 由引理 1 知, $2\dfrac{T_m}{T_i}, i = 1, 2, \cdots, m-1$ 相互独立且均服从 $\chi^2(2)$ 分布, 从而得证。

(2) 由式 (6) 及引理 1 知:

$$\begin{aligned} h(\beta) =& 2\sum_{i=1}^{m-1} \ln\left(\frac{\sum_{j=1}^{m-1}(r_j+1)Y_j + \left[n - \sum_{j=1}^{m-1}(r_j+1)\right]Y_m}{\sum_{j=1}^{i-1}(r_j+1)Y_j + \left[n - \sum_{j=1}^{i-1}(r_j+1)\right]Y_i}\right) = \\ & 2\sum_{i=1}^{m-1} \ln\left(1 + \frac{\sum_{j=i+1}^{m-1}(r_j+1)Y_j/Y_i + \left[n - \sum_{j=1}^{m}(r_j+1)\right]Y_m/Y_i}{\sum_{j=1}^{i-1}(r_j+1)Y_j/Y_i + \left[n - \sum_{j=1}^{i-1}(r_j+1)\right]}\right) = \\ & 2\sum_{i=1}^{m-1} \ln\left(1 + \frac{\sum_{j=i+1}^{m-1}(r_j+1)\mathrm{e}^{\beta(X_j-X_i)} + \left[n - \sum_{j=1}^{m}(r_j+1)\right]\mathrm{e}^{\beta(X_m-X_i)}}{\sum_{j=1}^{i-1}(r_j+1)\mathrm{e}^{\beta(X_j-X_i)} + \left[n - \sum_{j=1}^{i-1}(r_j+1)\right]}\right) \end{aligned} \tag{7}$$

式中: $e^{\beta(X_{i+1}-X_i)},\cdots,e^{\beta(X_m-X_i)}$ 为关于 β 的增函数; $e^{\beta(X_1-X_i)},\cdots,e^{\beta(X_{i-1}-X_i)}$ 为关于 β 的减函数。因此 $h(\beta)$ 在区间 $(-\infty,\infty)$ 为关于 β 的严格增函数。

(3) 类似于文献 [16] 中方程根的唯一性讨论, 由上述 $h(\beta)$ 表达式容易得出 $\lim\limits_{\beta\to-\infty}h(\beta)=-\infty$, $\lim\limits_{\beta\to+\infty}h(\beta)=+\infty$, 从而由零点定理知对任意 $t\in R$, 方程 $h(\beta)=t$ 在 $(-\infty,\infty)$ 有唯一根。

定理 设 $X_1<X_2<\cdots<X_m$ 是一个来自 Gompertz 分布的逐步增加定数截尾样本, 则:

(1) 形状参数 β 的置信水平为 $(1-\alpha)100\%$ 的置信区间为

$$\left[h^{-1}\left(\chi^2_{1-\alpha/2}(2m-2)\right),h^{-1}\left(\chi^2_{\alpha/2}(2m-2)\right)\right]$$

(2) θ,β 的置信水平为 $(1-\alpha)100\%$ 的联合置信域由下面不等式确定:

$$\left.\begin{aligned}&h^{-1}\left(\chi^2_{\frac{1+\sqrt{1+\alpha}}{2}}(2m-2)\right)\leqslant\beta\leqslant\\&\qquad h^{-1}\left(\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}(2m-2)\right)\\&\beta\chi^2_{\frac{1+\sqrt{1+\alpha}}{2}}(2m)(2T_m)^{-1}<\theta<\\&\qquad h^{-1}\left(\beta\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}(2m)\right)(2T_m)^{-1}\end{aligned}\right\}\tag{8}$$

式中: $\chi^2_\alpha(v)$ 表示自由度为 v 的 χ^2 分布的上侧 α 分位数; $h^{-1}(t)$ 为方程 $h(\beta)=t$ 的解。

证明 因为 $h(\beta)$ 是枢轴量且服从 $\chi^2(2m-2)$, 由引理 2 可得

$$\begin{aligned}1-\alpha=&P(\chi^2_{1-\frac{\alpha}{2}}(2m-2)<\\&h(\beta)<\chi^2_{\frac{\alpha}{2}}(2m-2))=\\&P(h^{-1}\chi^2_{1-\frac{\alpha}{2}}(2m-2))<\beta<\\&h^{-1}\chi^2_{\frac{\alpha}{2}}(2m-2)))\end{aligned}\tag{9}$$

式中, P 为事件发生的概率。

因 $h(\beta)$ 和 $\dfrac{2\theta T_m}{\beta}$ 均为枢轴量且相互独立, 且 $h(\beta)\sim\chi^2(2m-2)$, $\dfrac{2\theta T_m}{\beta}\sim\chi^2(2m)$, 于是

$$\begin{aligned}&P\Bigg(h^{-1}\left(\chi^2_{\frac{1+\sqrt{1+\alpha}}{2}}(2m-2)\right)<\beta<\\&h^{-1}\left(\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}(2m-2)\right),\\&\frac{\beta\chi^2_{\frac{1+\sqrt{1+\alpha}}{2}}(2m)}{2T_m}<\theta<\frac{\beta\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}(2m)}{2T_m}\Bigg)=\\&P\Bigg(\chi^2_{\frac{1+\sqrt{1+\alpha}}{2}}(2m-2)<h(\beta)<\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}(2m-2),\\&\chi^2_{\frac{1+\sqrt{1+\alpha}}{2}}(2m)<\frac{2\theta T_m}{\beta}<\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}(2m)\Bigg)=\\&P\left(\chi^2_{\frac{1+\sqrt{1+\alpha}}{2}}(2m-2)<h(\beta)<\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}(2m-2)\right)\\&P\left(\chi^2_{\frac{1+\sqrt{1+\alpha}}{2}}(2m)<\frac{2\theta T_m}{\beta}<\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}(2m)\right)=\\&\sqrt{1-\alpha}\sqrt{1-\alpha}=1-\alpha\end{aligned}$$

从而得证。

由上述讨论, 利用王炳兴 [17] 的逆矩估计方法可以推出 Gompertz 分布参数的逆矩估计。由引理 1 知 $-2i\ln(T_i/T_{i+1})\,(i=1,2,\cdots,m-1)$ 为准样本, 因此 β 参数的逆矩估计 $\hat\beta$ 为

$$h(\beta)=2m-4\tag{10}$$

由引理 2 知, 式 (10) 存在唯一解。

同样地, $Z_1,Z_2,\cdots,Z_m$ 为来自标准指数分布的准样本, 因此 θ 参数的逆矩估计 $\hat\theta$ 由下式确定:

$$\hat\theta=\frac{m\hat\beta}{\sum_{i=1}^m(r_i+1)e^{\hat\beta x_i}-n}\tag{11}$$

2 Monte Carlo 模拟

以上讨论了逐步增加定数截尾下 Gompertz 分布的点估计和区间估计, 按以下 Monte Carlo 方法可得 Gompertz 分布场合逐步增加定数截尾样本的随机数, 具体步骤 [18] 如下:

(1) 从标准指数分布产生 m 个相互独立的随机数 $Z_1,Z_2,\cdots,Z_m$;

(2) 令

$$\begin{aligned}Y_i=&\frac{Z_1}{n}+\frac{Z_2}{n-R_1-1}+\frac{Z_3}{n-R_1-R_2-2}+\cdots+\\&\frac{Z_i}{n-R_1-\cdots-R_{i-1}-i+1},\end{aligned}$$

式中: $i=1,2,\cdots,m$; $Y_1,Y_2,\cdots,Y_m$ 为逐步增加定数截尾下标准指数分布样本; n 为试验的产品总个数; $R_1,R_{2,}\cdots,R_{m-1}$ 为依次从试验中移离的产品个数。

(3) 最后, 令

$$X_i=\frac{1}{\beta}\ln\left(1+\frac{\beta Y_i}{\theta}\right),\ i=1,2,\cdots,m,$$

式中: $X_1,X_2,\cdots,X_m$ 为来自 Gompertz 分布的逐步增加定数截尾样本。

为了考察逆矩估计的精度, 取 Gompertz 分布参数 $\theta = 1.5, \beta = 3$, 从该分布产生 m 个逐步增加定数截尾样本 $x_1, x_2, \cdots, x_m$。每种试验方案产生 100 组模拟样本并且计算逆矩估计的相对偏差与均方误差, 具体试验安排与估计结果如表 1 所示。

表 1 为逐步增加定数截尾下 Gompertz 分布的试验方案和估计结果, 由估计的相对偏差 (R_{Bias}) 和相对均方误差 (R_{MSE}) 可以看出, 逆矩估计的精度较高, 说明该估计有效。

3　实例研究

例 1　利用 Wu 等 [3] 给出的无瘤时间的逐步增加定数截尾样本证实本文所提方法, 令 $n = 30, m = 16$, 具体试验方案与观测数据如表 2 所示。

利用式 (10)、(11) 分别得到参数 θ 和 β 的逆矩估计分别为 $\hat{\theta} = 0.051\ 1$ 和 $\hat{\beta} = 0.000\,23$。为了得到的 95% 置信区间, 需要 2 个分位数: $\chi^2_{0.025}(30) = 46.979\ 2$ 和 $\chi^2_{0.025}(30) = 16.790\ 8$。利用定理得到 β 的 95% 置信区间为 $(0.045\ 9, 0.145\ 8)$, 该结果跟文献 [3] 比较接近, 说明了方法的可行性。

进一步, 为获得 θ 和 β 的 95% 联合置信区间, 需要分位数: $\chi^2_{0.611\ 8}(30) = 27.218\ 0$, $\chi^2_{0.388\ 2}(30) = 31.561\ 5$, $\chi^2_{0.611\ 8}(32) = 29.144\ 8$ 和 $\chi^2_{0.388\ 2}(32) = 33.634\ 3$, θ 和 β 的 95% 联合置信区间由下列不等式确定:

$$\begin{cases} 0.007\ 8 < \beta < 0.169\ 8, \\ \dfrac{14.572\ 4\beta}{T_{16}} < \theta < \dfrac{16.817\ 1\beta}{T_{16}} \end{cases}$$

表 1　逐步增加定数截尾下 Gompertz 分布的试验方案与估计结果

Tab. 1　Test scheme and estimate results of Gompertz distribution under progressive type II censoring

方案	n, m	$R_1, \cdots, R_{m-1}, R_m$	$R_{\text{Bias}}(\hat{\theta})$	$R_{\text{MSE}}(\hat{\theta})$	$R_{\text{Bias}}(\hat{\beta})$	$R_{\text{MSE}}(\hat{\beta})$
1	30, 9	4, 4, 4, 4, 0, 0, 0, 0, 5	0.016 8	4.125 6e^{-4}	0.657 6	0.786 5
2	40, 10	5, 5, 5, 5, 0, 0, 0, 0, 0, 10	0.008 9	2.448 7e^{-4}	0.487 5	0.608 7
3	50, 10	6, 6, 6, 6, 0, 0, 0, 0, 0, 16	0.007 8	1.542 1e^{-4}	0.643 5	0.713 2
4	60, 10	7, 7, 7, 7, 0, 0, 0, 0, 0, 12	0.003 7	1.456 9e^{-4}	0.476 9	0.589 5

表 2　基于无瘤时间的逐步增加定数截尾样本

Tab. 2　Progressive type II censored sample based on the tumor-free data

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
x_i	60	63	63	63	66	68	70	77	84	91	91	94	101	109	112	115
R_i	1	0	0	2	1	0	1	0	2	0	2	0	2	0	0	3

参考文献:

[1] 周会会. Gompertz 分布尺度参数的最短区间估计 [J]. 山东理工大学学报 (自然科学版), 2020, 34(3): 70-73.

[2] WU J W, HUNG W L, TSAI C H. Estimation of the parameters of the Gompertz distribution under the first failure-censored sampling plan [J]. Statistics, 2003, 37(6): 517-525.

[3] WU S J, CHANG T, TSAI T R. Point and interval estimations for the Gompertz distribution under progressive type-II censoring [J]. METRON-International Journal of Statistics, 2003, LXI(3): 403-418.

[4] LENART A. The moments of the Gompertz distribution and maximum likelihood estimation of its parameters [J]. Scandinavian Actuarial Journal, 2014, 2014(3): 255-277.

[5] SANKU D, FERNANDO A M, DEVENDRA K. Statistical properties and different methods of estimation of Gompertz distribution with application [J]. Journal of Statistics and Management Systems, 2018, 21(5): 839-876.

[6] HANAA H A, ANHAR S A. Different method estimations of the three parameters of exponentiated Gompertz distribution [J]. Applied Mathematics and Information Sciences, 2016, 10(2): 705-710.

[7] ELIWA M S, EI-MORSHEDY M, IBRAHIM M. Inverse Gompertz distribution: Properties and different estimation methods with application to complete and censored data

[J]. Annals of Data Science, 2019, 6: 321-339.

[8] EGHWERIDO J T, NZEI L C, AGU F I. The alpha power Gompertz distribution: Characterization, properties and applications [J/OL]. The Indian Jouranl of Statistics, https://doi.org/10.1007/s13171-020-00198-0.

[9] JOSMAR M, ANDRÉ F M, SANKU D. Unit-Gompertz distribution with applications [J]. Statistica, 2019, 79(1): 25-43.

[10] ISMIAL A. On the optimal design of step-stress partially accelerated life tests for the Gompertz distribution with type I censoring [J/OL]. https://www.researchgate.net/publication/228625178.

[11] SARAÇOĞLU B, KAYA M F. Maximum likelihood estimation and confidence intervals of system reliability for Gompertz distribution in stress-strength models [J]. Secuk Journal of Applied Mathematics, 2007, 8(2): 25-36.

[12] 王蓉华, 徐晓岭, 施宏伟. Gompertz 分布 TFR 模型多步步进应力加速寿命试验的统计分析 [J]. 应用概率统计, 2009, 25(1): 47-59.

[13] ESSAM A A. Estimation of some lifetime parameters of generalized Gompertz distribution under progressively tyep-II censored data [J]. Appied Mathematical Modelling, 2015, 39(18): 5567-5578.

[14] THOMAS D R, WILSON W M. Linear order statistic estimation for the two parameter Weibull and extreme-value distribution under Type-II progressively censored samples [J]. Technometrics, 1972, 14: 679-691.

[15] 王炳兴. Burr Type XII 分布的统计推断 [J]. 数学物理学报, 2008, 28A(6): 1103-1108.

[16] 张志华. 加速寿命试验及其统计分析 [M]. 北京: 北京工业大学出版社, 2002.

[17] 王炳兴. Weibull 分布的统计推断 [J]. 应用概率统计, 1992(4): 357-364.

[18] 马永传, 武东, 朱雅敏. 广义逐次定数截尾下 Pareto 分布的统计分析 [J]. 通化师范学院学报, 2016, 37(10): 22-24.