行政院國家科學委員會補助專題研究計畫成果報告

在逐步型II設限下有關Gompertz分配參數之估計

Estimation in the Gompertz Distribution base on Progressive Type II Censoring

計畫類別：個別型計畫
計畫編號：NSC 91-2118-M-032-011
執行期間：91年8月1日至92年7月31日

計畫主持人：吳碩傑
共同主持人：張春桃
計畫參與人員：林英博

成果報告類型（依經費核定清單規定繳交）：精簡報告

本成果報告包括以下應繳交之附件：無

處理方式：立即公開查詢

執行單位：淡江大學統計學系

中　華　民　國　九十二　年　八　月　十九　日

在逐步型II設限下有關Gompertz分配參數之估計

吳碩傑

中文摘要

在壽命試驗中，實驗者經常不會等所有測試單位的壽命都被觀察到才停止試驗，而是利用設限抽樣的方式來觀察到一部分測試單位的壽命便停止試驗。本研究考慮一種多用途的逐步型 II 設限方案，探討在此設限方案下 Gompertz 分配參數的估計問題。我們將利用最大概似法來求出 Gompertz 分配參數的估計式，並利用樞紐量方法來建立 Gompertz 分配參數的信賴區間及聯合信賴區域。此外，我們也將舉一個實際的例子來闡述我們所得到的結果。

關鍵詞：信賴區間；聯合信賴區域；最大概似估計式；樞紐量；逐步型 II 設限。

Abstract

We obtain estimation results concerning a progressively type-II censored sample from the Gompertz distribution. The maximum likelihood method is used to derived the point estimators of the parameters of the Gompertz distribution. An exact confidence interval and an exact joint confidence region for the parameters of the Gompertz distribution are constructed. Two numerical examples are presented to illustrate the methods of inference developed here.

Keywords: Confidence interval; Joint confidence region; Maximum likelihood estimator; Pivot; Progressively type-II censoring.

1 Introduction

The Gompertz distribution was originally introduced by Gompertz (1825). This distribution is widely used to describe human mortality and establish actuarial tables. The probability density function of a Gompertz distribution has the form:

$$f(x) = \begin{cases} \lambda e^{cx} \exp\left\{-\frac{\lambda}{c}(e^{cx}-1)\right\}, & \text{for } x > 0, \\ 0, & \text{elsewhere,} \end{cases} \tag{1}$$

where $c > 0$ and $\lambda > 0$ are the parameters. In recent years, this distribution has been studied by several authors. Garg *et al.* (1970) investigated the properties of the Gompertz distribution and derived the maximum likelihood estimators of the parameters. Gordon (1990) examined the feasibility of maximum likelihood estimation of a mixture of two Gompertz distributions when censoring occurs. Chen (1997) developed an exact confidence interval and a joint confidence region for the parameters of Gompertz distribution.

In this study, we consider a censoring scheme called progressive type-II censoring. Under this scheme, n units are placed on a test at time zero, and m failures are going to be observed. When the first failure is observed, r_1 of the surviving units are randomly selected and removed. At the second observed failure, r_2 of the surviving units are randomly selected and removed. This experiment stops at the time when the m-th failure is observed and the remaining $r_m = n - r_1 - r_2 - \cdots - r_{m-1} - m$ surviving units are all removed. The statistical inference on the parameters of some distributions under progressive type-II censoring has been investigated by several authors such as Balakrishnan and Aggarwala (2000).

In this article, we consider progressive type-II censored data from a Gompertz distribution. We obtain the maximum likelihood estimators of the parameters in Section 2. We derive an exact confidence interval for the parameter c and an exact joint confidence region for the parameters c and λ in Section 3. A numerical example is presented for illustration in Section 4. Some conclusions are made in Section 5.

2 Point Estimation of Parameters

In this section, the maximum likelihood estimators (MLEs) for the parameters of Gompertz distribution based on progressive type-II censoring are derived.

Let $X_1, X_2, \ldots, X_m$ be a progressively type-II censored sample from the Gompertz distribution, with censoring scheme $\boldsymbol{r} = (r_1, r_2, \ldots, r_m)$. The log-likelihood function based on the progressively type-II censored sample is given by

$$\log L(c, \lambda) \propto m \log \lambda + c \sum_{i=1}^{m} x_i - \frac{\lambda}{c}\left[\sum_{i=1}^{m}(r_i+1)e^{cx_i} - n\right],$$

and thus we have the likelihood equations for λ and c to be

$$\frac{\partial \log L}{\partial \lambda} = \frac{m}{\lambda} - \frac{1}{c}\left[\sum_{i=1}^{m}(r_i+1)e^{cx_i} - n\right] = 0, \tag{2}$$

and

$$\frac{\partial \log L}{\partial c} = \sum_{i=1}^{m} x_i - \frac{\lambda}{c}\sum_{i=1}^{m}(r_i+1)x_i e^{cx_i} + \frac{\lambda}{c^2}\left[\sum_{i=1}^{m}(r_i+1)e^{cx_i} - n\right] = 0. \tag{3}$$

The MLEs $\hat{c}$ and $\hat{\lambda}$ can be obtained by solving the likelihood equations. Equation (2) yields the MLE of λ to be

$$\hat{\lambda} = \frac{m\hat{c}}{\sum_{i=1}^{m}(r_i+1)e^{\hat{c}x_i} - n}. \tag{4}$$

Equation (3), in conjunction with the MLE of λ in (4), reduces to

$$\sum_{i=1}^{m} x_i - \frac{m\sum_{i=1}^{m}(r_i+1)x_i e^{\hat{c}x_i}}{\sum_{i=1}^{m}(r_i+1)e^{\hat{c}x_i} - n} + \frac{m}{\hat{c}} = 0. \tag{5}$$

Since (5) cannot be solved analytically for $\hat{c}$, some numerical methods such as Newton's method must be employed.

3 Interval Estimation of Parameters

In this section, an exact confidence interval for c and an exact joint confidence region for c and λ are discussed. Let $Y_i = \frac{\lambda}{c}(e^{cX_i} - 1)$, $i = 1, 2, \ldots, m$. It is easy to show that $Y_1 < Y_2 < \cdots < Y_m$ is a progressively type-II censored sample from an exponential distribution with mean 1. Let us consider the transformation $Z_1 = nY_1$ and $Z_i = (n - r_1 - \cdots - r_{i-1} - i + 1)(Y_i - Y_{i-1})$, for $i = 2, \ldots, m$. Thomas and Wilson (1972) established that the generalized spacings $Z_1, Z_2, \ldots, Z_m$ are independent and identically distributed as an exponential distribution with mean 1. Hence,

$$V = 2Z_1 = 2nY_1$$

has a chi-square distribution with 2 degrees of freedom and

$$U = 2\sum_{i=2}^{m} Z_i = 2\left[\sum_{i=1}^{m}(r_i+1)Y_i - nY_1\right]$$

has a chi-square distribution with $2m-2$ degrees of freedom. We can also find that U and V are independent random variables. Let

$$T_1 = \frac{U}{(m-1)V} = \frac{\sum_{i=1}^{m}(r_i+1)Y_i - nY_1}{n(m-1)Y_1}, \tag{6}$$

and

$$T_2 = U + V = 2\sum_{i=1}^{m}(r_i+1)Y_i. \tag{7}$$

It can be seen that T_1 has an F distribution with $2m-2$ and 2 degrees of freedom and T_2 has a chi-square distribution with $2m$ degrees of freedom. Furthermore, by Johnson *et al.* (1994, p.350), T_1 and T_2 are independent.

To obtain the confidence interval for c and the joint confidence region for c and λ, we need the following lemmas.

Lemma 1. *For any* $0 < a < b$,

$$g(c) = \frac{e^{cb}-1}{e^{ca}-1}$$

is a strictly increasing function of c *for any* $c \neq 0$.

Proof. The proof of Lemma 1 can be found in Li (2000). □

Lemma 2. *Suppose that* $0 < a_1 < a_2 < \cdots < a_m$. *Let*

$$T_1(c) = \frac{\sum_{i=1}^{m}(r_i+1)(e^{ca_i}-1) - n(e^{ca_1}-1)}{n(m-1)(e^{ca_1}-1)}.$$

Then, $T_1(c)$ *is strictly increasing in* c *for any* $c \neq 0$. *Furthermore, if*

$$t \neq \frac{\sum_{i=1}^{m}(r_i+1)a_i - na_1}{n(m-1)a_1},$$

the equation $T_1(c) = t$ *has a unique solution for any* $c \neq 0$.

Proof. The proof of Lemma 2 can be found in Wu (2003). □

Let $F_{\alpha(\nu_1,\nu_2)}$ denote the percentile of F distribution with right-tail probability α and ν_1 and ν_2 degrees of freedom. The following theorem gives an exact confidence interval for the parameter c.

Theorem 1. *Suppose that X_i, $i = 1, \ldots, m$, are the order statistics of a progressively type-II censored sample from a sample of size n from a distribution which has density function (1), with censoring scheme $(r_1, r_2, \ldots, r_m)$. Then a $100(1-\alpha)\%$ confidence interval for c is:*

$$\left(\varphi(X_1, \ldots, X_m, F_{1-\frac{\alpha}{2}(2m-2,2)}) \quad , \quad \varphi(X_1, \ldots, X_m, F_{\frac{\alpha}{2}(2m-2,2)})\right),$$

where $0 < \alpha < 1$ and $\varphi(X_1, \ldots, X_m, t)$ is the solution of c for the equation

$$\frac{\sum_{i=1}^{m}(r_i+1)e^{cX_i} - ne^{cX_1}}{n(m-1)(e^{cX_1}-1)} = t.$$

Proof. The proof of Theorem 1 can be found in Wu (2003). □

Note that the lower confidence bound for c obtained in Theorem 1 may be less than 0. Since c is restricted to be nonnegative, the appropriate procedure is to replace any negative bound with 0.

Let $\chi^2_{\alpha(\nu)}$ be the percentile of chi-square distribution with right-tail probability α and ν degrees of freedom. An exact joint confidence region for the parameters c and λ are given in the following theorem.

Theorem 2. *Suppose that X_i, $i = 1, \ldots, m$, are the order statistics of a progressively type-II censored sample from a sample of size n from a distribution which has density function (1), with censoring scheme $(r_1, r_2, \ldots, r_m)$. Then a $100(1-\alpha)\%$ joint confidence region for c and λ is determined by the following inequalities:*

$$\begin{cases} \varphi\left(X_1, \ldots, X_m, F_{\frac{1+\sqrt{1-\alpha}}{2}(2m-2,2)}\right) < c < \varphi\left(X_1, \ldots, X_m, F_{\frac{1-\sqrt{1-\alpha}}{2}(2m-2,2)}\right) \\ \dfrac{c\chi^2_{\frac{1+\sqrt{1-\alpha}}{2}(2m)}}{2\sum_{i=1}^{m}(r_i+1)(e^{cX_i}-1)} < \lambda < \dfrac{c\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}(2m)}}{2\sum_{i=1}^{m}(r_i+1)(e^{cX_i}-1)}, \end{cases}$$

where $0 < \alpha < 1$ and $\varphi(X_1, \ldots, X_m, t)$ is the solution of c for the equation

$$\frac{\sum_{i=1}^{m}(r_i+1)e^{cX_i} - ne^{cX_1}}{n(m-1)(e^{cX_1}-1)} = t.$$

Proof. The proof of Theorem 2 can be found in Wu (2003). □

4 Illustrative Examples

To illustrate the use of the estimation methods discussed in the article, the following example is discussed.

Table 1: Progressively type-II censored sample generated from the tumor-free time 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

Example 1. Lee (1992, table 3.4) presented data on the tumor-free time in days of 30 rats fed with unsaturated diet. In analyzing these data, Chen (1997) assumed a Gompertz distribution for the tumor-free times. For the purposes of illustrating the methods discussed here, a progressively type-II censored sample of size $m = 16$ was randomly selected from the $n = 30$ observations presented in Lee's table 3.4. The observations and censoring scheme are reported in Table 1. The 16 observations were generated in the following manner. We selected the smallest observation from the original 30 observations, and then $r_1 = 1$ of remaining 29 is selected at random and removed. We then selected the next 3 ordered observations and then removed $r_4 = 2$ at random from the remaining 25. The process is continued according to the censoring scheme $\boldsymbol{r}$ to produce this data set.

Using the iterative formula described in Section 2, we determine the MLEs of c and λ to be $\hat{c} = 0.0505$ and $\hat{\lambda} = 0.00024$, respectively. To find a 95% confidence interval for c, we need the percentiles $F_{0.025(30,2)} = 39.4646$ and $F_{0.975(30,2)} = 0.2391$. By Theorem 1 and using the Fortran IMSL nonlinear equation solver, the 95% confidence interval for c is $(0.0445, 0.1464)$.

Furthermore, to obtain a 95% joint confidence region for c and λ, we need the percentiles $F_{0.0127(30,2)} = 78.4528$, $F_{0.9873(30,2)} = 0.1972$, $\chi^2_{0.0127(32)} = 52.4848$ and $\chi^2_{0.9873(32)} = 16.8214$. By Theorem 2, a 95% joint confidence region for c and λ is determined by the following inequalities:

$$\begin{cases} 0.0405 < c < 0.1595 \\ \dfrac{16.8214c}{2\sum_{i=1}^{16}(r_i+1)(e^{cx_i}-1)} < \lambda < \dfrac{52.4848c}{2\sum_{i=1}^{16}(r_i+1)(e^{cx_i}-1)} \end{cases}.$$

□

5 Conclusions

We use the maximum likelihood method to obtain the point estimators of the parameters of a Gompertz distribution based on a progressively Type II censored sample. We also construct an exact confidence interval and an exact joint confidence region for the parameters, respectively. To illustrate the use of proposed point and interval estimators, two numerical data sets are analyzed.

References

Balakrishnan, N. & Aggarwala, R. (2000). *Progressive Censoring – Theory, Methods, and Applications.* Boston: Birkhäuser.

Chen, Z. (1997). Parameter estimation of the Gompertz population. *Biometrical Journal* **39**, 117-124.

Garg, M. L., Rao, B. R. & Redmond, C. K. (1970). Maximum-likelihood estimation of the parameters of the Gompertz survival function. *Applied Statistics* **19**, 152-159.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality and on the new mode of determining the value of life contingencies. *Phil. Trans. R. Soc.* **A 115**, 513-580.

Gordon, N. H. (1990). Maximum likelihood estimation for mixtures of two Gompertz distributions when censoring occurs. *Communications in Statistics – Simulation and Computation* **19**, 733-747.

Johnson, N. L., Kotz, S. & Balakrishnan, N. (1994). *Continuous Univariate Distributions, Volume 1*, 2nd edition. New York: Wiley.

Lee, E. T. (1992). *Statistical Methods for Survival Data Analysis*, 2nd edition. New York: Wiley.

Li, P.-L. (2000). Interval estimation of the parameters of Gompertz and extreme value distributions based on the doubly type II censored samples. Master thesis, Tamkang University, Department of Statistics (in Chinese).

Thomas, D. R. & Wilson, W. M. (1972). Linear order statistic estimation for the two-parameter Weibull and extreme value distributions from type-II progressively censored samples. *Technometrics* **14**, 679-691.

Wu, S.-J. (2003). Point and interval estimations for the Gompertz distribution under progressive type-II censoring. Submitted.