行政院國家科學委員會補助專題研究計畫成果報告

※※※※※※※※※※※※※※※※※※※※※※※※※※※
※　　　　　　　　　　　　　　　　　　　　　※
※在第一失敗-設限抽樣方案下 GOMPERTZ 分配的參數估計(2/2) ※
※　　　　　　　　　　　　　　　　　　　　　※
※　　　　　　　　　　　　　　　　　　　　　※
※※※※※※※※※※※※※※※※※※※※※※※※※※※

計畫類別：☑個別型計畫　　□整合型計畫

計畫編號：NSC 89-2118-M-032-013

執行期間：88 年 8 月 1 日至 89 年 7 月 31 日

計畫主持人：吳忠武

計畫參與人員：　蔡志輝（研究生）

執行單位：　淡 江 大 學

中　華　民　國　89　年　6　月　30　日

行政院國家科學委員會專題研究計劃成果報告

在第一失敗-設限抽樣方案下 GOMPERTZ 分配的參數估計(2/2)

計劃編號：NSC 89-2118-M-032-013
執行期限：1999 年 8 月 1 日至 2000 年 7 月 31 日
主持人：吳忠武　研究生：蔡志輝　執行機構及單位名稱：淡江大學統計學系

一、中文摘要

我們在探討一產品或生物之壽命時，常會使用一些分配之圖形曲線來描述其壽命。本篇報告主要探討在第一失敗-設限抽樣方案下 Gompertz 分配之參數估計，而此分配最早由 Gompertz (1825)提出，用來描述人類死亡率和建立精算表。本篇報告大致可區分成為兩部份：第一部份為在第一失敗-設限抽樣方案下 Gompertz 分配參數之區間估計；第二部份為 Gompertz 分配參數之點估計。

Abstract

When we discuss the lifetime of a produce or a creature, we usually use some curves of distributions to fix the lifetime. This paper aims to provide the parameter estimates of the Gompertz distribution under the first failure-censored sampling plans. The Gompertz distribution was introduced by Gompertz (1825) to describe human mortality and establish actuarial tables.

This paper is divided into two main parts. The first part of the present paper deals with the interval estimation under the first failure-censored sampling plans, while the second part handles the point estimation.

Keywords: Gompertz Distribution, Joint Confidence Region, Exact Estimation, Order Statistics, First Failure-censored.

二、緣由與目的

首先於第一部份，Chen(1997)曾提出在型 II 設限(censored)資料下求出 Gompertz 分配參數的信賴區間(confidence interval)及聯合信賴區域(joint confidence region)。而此篇報告我們考慮的為第一失敗-設限抽樣方案，即在 m 條生產線上各抽出 n 個產品樣本，觀察並記錄其產品壽命，在每一條生產線上，我們只取第一個壞掉之產品壽命，共可得到 m 個產品之壽命資料。此種抽樣方法將可為公司企業節省大量的金錢與寶貴的時間。Balasooriya (1995)也曾提到在此抽樣方法下有關指數分配之研究。而在此，我們假設原先的所有樣本皆來自 $Gompertz(c,\lambda)$ 分配，則利用此抽樣方法最後所得到的 m 個樣本資料 $X_{(1)1}, X_{(1)2}, \cdots, X_{(1)m}$ 會來自 $Gompertz$ $(c, \lambda^* = n\lambda)$，而 $X'_{(1)} < X'_{(2)} < \cdots < X'_{(m)}$ 為其對應的有序統計量(order statistics)，我們使用變數變換 $Y_{(i)} = \dfrac{\lambda^*}{c}(e^{cX'_{(i)}} - 1)$，則可以得到以下隨機變數：$U = 2(\sum_{i=1}^{m} Y_{(i)} - mY_{(1)}) \sim \chi^2(2m-2)$ 和 $V = 2mY_{(1)} \sim \chi^2(2)$，並且再透過下面的變數轉換 $\xi = \dfrac{U}{(m-1)V} \sim F(2m-2, 2)$ 和 $\zeta = (U+V) \sim \chi^2(2m)$，利用 ξ 與 ζ 我們就可以求出參數 c 之區間估計，及 c 與 λ 之聯合信賴區域。我們將舉例說明，並且最後利用模擬估計區間估計中的顯著水準 α。

接著於第二部份，Garg, Rao 和 Damaraju

(1970)曾討論 Gompertz 分配的性質且得到參數的最大概似估計量。Gordon (1990)則提供兩個 Gompertz 分配混合的最大概似估計法。Pan 和 Chen (1998)提出在不同機率分配下可靠度壽命估計方法之比較。Franses (1994)提出如何利用 Gompertz 曲線來描述某產品的壽命。而此篇報告所要研究的是有關 Gompertz 分配中參數c和λ的點估計，而 Garg, Rao 和 Redmond (1970)已提出最大概似法來估計此二參數，在此我們將拿來做比較。而本報告在此要提出一個非線性的最小平方法來估計參數c和λ，其步驟有三：

(a)利用 Gompertz 分配之累積密度函數

$$F(x)=1-\exp[-\frac{\lambda}{c}(e^{cx}-1)] \tag{1.1}$$

，經二次的對數轉換後形成

$$\ln\ln(\frac{1}{1-F(x)})=\ln\lambda+\ln(\frac{e^{cx}-1}{c}) \tag{1.2}$$

的線性關係。

(b)假設有一組大小為 n 的隨機樣本 $X_1, X_2, \cdots, X_n$ 來自 $Gompertz(c, \lambda)$ 具有累積密度函數如(1.1)，而 $X_{(1)}<X_{(2)}<\cdots<X_{(n)}$ 為其對應的有序統計量，則可得

$$E[F(X_{(i)})]=\frac{i}{n+1}, i=1,2,\cdots,n \tag{1.3}$$

所以用(1.3)中的 $\frac{i}{n+1}$ 來估計(1.2)中的 $F(x)$。

(c)利用 $\min\sum_{i=1}^{n}[\ln\ln\frac{n+1}{n-i+1}-\ln\lambda-\ln(\frac{e^{cx_{(i)}}-1}{c})]^2$ 來估計c和λ，而且本研究也利用另二個估計 $\frac{i-0.5}{n}$ 或 $\frac{i-0.3}{n+0.4}$，(參考 Kapur 和 Lamberson(1977)、Ross(1987))來估計(1.2)中的$F(x)$，然後再利用步驟(c)來估計c和λ。

另外，也考慮若$1-F(x)$是來自於均勻分配$U(0,1)$，則$-\ln[1-F(x)]$是來自於指數分配$E(1)$。因此，可利用下面的結果 $E[-\ln(1-F(x_{(i)}))]=\sum_{j=1}^{i}\frac{1}{n-j+1}$(也可參考 Lawless(1982))和(1.2)提出另一個非線性最小平方估計量來估計c和λ。

其次，本研究我們也考慮在第一失敗-設限抽樣方案下所產生的樣本資料 $X_{(1)1}, X_{(1)2}, \cdots, X_{(1)m}$ 會來自於 Gompertz (c, $\lambda^*=n\lambda$) 分配，而 $X'_{(1)}<X'_{(2)}<\cdots<X'_{(m)}$ 為其對應的有序統計量，因此(1.2)和(1.3)也分別改寫成為

$$\ln\ln(\frac{1}{1-F(x)})=\ln n+\ln\lambda+\ln(\frac{e^{cx}-1}{c}) \tag{1.4}$$

和

$$E[F(X_{(i)})]=\frac{i}{m+1}, i=1,2,\cdots,m \tag{1.5}$$

所以以(1.5)中 $\frac{i}{m+1}$ 來估計(1.4)中的$F(x)$，並且利用

$$\min\sum_{i=1}^{m}[\ln\ln\frac{m+1}{m-i+1}-\ln n-\ln\lambda-\ln(\frac{e^{cx_{(i)}}-1}{c})]^2 \tag{1.6}$$

來估計c和λ。同理，也可利用$\frac{i-0.5}{m}$或$\frac{i-0.3}{m+0.4}$來估計(1.4)中的$F(x)$，然後再利用非線性最小平方法及(1.4)來估計c和λ。

最後，我們也做一些蒙地卡羅(Monte Carlo)模擬來評估這些新的非線性最小平方估計量與 Garg, Rao 和 Redmond(1970)提出的最大概似估計量之好壞。

三、結果與討論

從模擬結果我們可以得知，在區間估計部分，當生產線(m)或每一生產線之產品樣本數(n)增加時,其模擬出之顯著水準與真值

皆差不多，即$\hat{\alpha} \cong \alpha$。而且標準誤隨著 m 或 n 增加而減少，顯示我們估計顯著水準是非常精確。

至於點估計的部分：在完整資料下之點估計，我們可以從模擬結果中發現，其中最小平方法所估計出之標準誤的值普遍皆會比最大概似法所估計出之標準誤的值來得較小。另外，當我們抽出產品樣本數為 30 個(n=30)或 10 個(n=10)時，其估計結果差不多，而若我們考慮成本因素時，那我們當然會選擇以較少樣本數(n=10)之條件下來做估計。總體來講，在完整資料下最小平方法估計之結果普遍皆比最大概似法估計之結果來得好，而最小平方法中又以有加權的情況下較好。

在第一失敗-設限抽樣方案下之點估計，當我們固定各生產線抽出之產品樣本數(如：n=30,n=10)時，則較多生產線(m=30)會比較少生產線(m=10)之估計結果來得好。而若當我們固定生產線數(如：m=30,m=10)時，則抽出較少產品樣本數(n=10)反而會比抽出較多產品樣本數(n=30)之估計結果來得好。以表說明如下：

樣本數\生產線數	m=10	m=30
n=10	***	****
n=30	*	**

[註]****表示估計得最好,而*表示估計得最差。

而從模擬中我們也可以很容易地發現，其中最小平方法所估計出之標準誤的值，普遍皆會比最大概似法所估計出之標準誤的值來得較小。總體來講，在第一失敗-設限抽樣方案下，最小平方法估計之結果普遍皆比最大概似法估計之結果來得好，而最小平方法中又以有加權的情況下較好。

四、計劃成果自評

我們使用第一失敗-設限抽樣方案，對 Gompertz 分配的參數估計上，由於可以較少的樣本達到有效的估計，在應用上將可以作為有效節省成本的一種方案。另外，未來的研究方向，也可以朝向不同型態的設限資料或抽樣方法，做更進一步的發展。

五、參考文獻

[1]Balasooriya, U. (1995), Failure-Censored Reliability Sampling plans for the Exponential Distribution. *J. Statist. Comput. Simul.* **52**, 337-349.

[2]Barnett, V. (1975), Probability Plotting Methods and Order Statistics. *Appl. Statist.* **24**, No 1, 95.

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

[4]D'Agostino, R. B. and Stephens, M. A. (1986), *Goodness-of-fit Tech-niques*, Marcel Dekker, Inc., New York.

[5]Faucher, B. and Tyson, W. R. (1988), On the determination of Weibull Parameters. *Journal of Materials Science Letters* **7**, 1199-1203.

[6]Franses, P. H. (1994), Fitting a Gompertz curve. *Journal of the Operation Research Society* **45**, 109-113.

[7]Garg, M. L., Rao, B. R. and Redmond, C. K. (1970), Maximum likelihood estimation of the parameters of the Gompertz survival function. *Journal of the Royal Statistical Society C* **19**, 152-159.

[8]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**,733-747.

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

[10]IMSL User's Manual Math / Library (1989), IMSL , Int.,Houston, Texas.

[11]Kapur, K. C. and Lamberson, L. R. (1977),

Reliability in engineering design. Wiley, New York.

[12]Langlois, R. (1991), Estimation of Weibull parameters. *Journal of Materials Science Letters* **10**,1049-1051.

[13]Lawless, J. F. (1982), *Statistical Models and Methods for Lifetime Data.*

[14]Makany, R. (1991), A theoretical basis of Gompertz's curve. *Biometrical Journal* **33**, 121-128.

[15]Pan, J. N. and Chen, M. B. (1998), Reliability Estimation Methods under Various Probability Distribution Functions. *J. Chin. Stat. Association* **35(2)**,173-190.

[16]Rao, B. R. and Damaraju, C. V. (1992), New better than used and other concepts for a class of life distribution. *Biometrical Journal* **34**, 919-935.

[17]Read, C. B. (1983), *Gompertz distribution. Encyclopedia of Statistical Sciences.* Wiley, New York.

[18]Ross, S. M. (1987), *Introduction to probability and statistics for engineers and scientists.* Wiley, New York.