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 Type I (MINUMUM) EXTREME VALUE 和 WEIBULL 分配的參數估計

計畫類別：☑個別型計畫　　□整合型計畫
計畫編號：NSC 89-2118-M-032-002
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： 林尚賢（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 Type I (MINUMUM) EXTREME VALUE 和 WEIBULL 分配的參數估計

計劃編號：NSC 89-2118-M-032-002
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：林尚賢　執行機構及單位名稱：淡江大學統計學系

一、中文摘要

在許多自然科學及工業生產的研究中，生物或是產品的壽命往往是研究的重點，而通常在生物或是產品的壽命檢測上，需要花費許多的時間、人力和金錢。本篇報告所討論的主題，是在第一失敗-設限抽樣方案(first failure-censored sampling plans)下，對 Weibull 分配和 Extreme-value 分配的參數進行估計，大致上可分為兩大部分進行討論：第一部份為在第一失敗-設限抽樣方案下參數的區間估計；第二部分為參數的點估計部分。

Abstract

In many researches related to nature science or industrial production, the main issue is often center on the life span of organism or product. None the less, the inspection of life span usually costs a lot, be it time, money or human effort. This paper aims to provide the parameter estimates of the Weibull distribution and the Extreme-value distribution under the first failure-censored sampling plans. 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: Weibull Distribution, Extreme-value Distribution, Joint Confidence Region, Order Statistics, First Failure-censored, Point Estimation.

二、緣由與目的

首先於第一部份，Chen(1998)曾提出在型Ⅱ設限(type Ⅱ censored)資料下 Weibull 和 Extreme-value 分配參數的信賴區間(confidence interval)及聯合信賴區域(joint confidence region)。而我們在本篇報告中則使用第一失敗-設限抽樣方案(first failure-censored sampling plans)，也就是在 m 組樣本中分別抽出 n 個樣本紀錄其壽命，而每一組樣本我們只取第一個失敗的樣本壽命，最後，我們可以得到 m 個樣本的壽命資料。如此一來，將此法應用在工業生產或生物統計的研究上，可以節省許多時間和金錢。

我們以此抽樣方案所得的 m 個 Weibull 分配樣本資料 $X_{(1)1}, X_{(1)2}, \mathrm{K}, X_{(1)m}$ 將其排序使得 $X'_{(1)1} \le X'_{(1)2} \le \mathrm{K} \le X'_{(1)m}$ 為對應之順序統計量(order statistic)，並且利用變數變換 $Y_{(i)} = n\left(\frac{X'_{(1)i}}{\eta}\right)^{\beta}$ 得到隨機變數 $U = 2mY_{(1)} \sim \chi^2(2)$，$V = 2\left(\sum_{i=1}^{m} Y_{(i)} - mY_{(1)}\right) \sim \chi^2(2m-2)$，並且利用 $\zeta' = \frac{V}{(m-1)U} \sim F(2m-2,2)$ 和 $\delta' = U + V \sim \chi^2(2m)$ 求得參數 β 之區間估計與參數 β 和 η 的聯合信賴區域。相同的方法我們也使用在 Extreme-value 分配的參數 θ 和 ξ 估計上，利

用變數變換 $Z_i = n\exp(\frac{X'_{(1)i}-\xi}{\theta})$ 進行與 Weibull 分配相同的推論。

接著，在第二部分中我們引用 Bain 和 Engelhardt(1991) 提出 Weibull 分配與 Extreme-value 分配的最大概似估計量，並且我們也將提出一些具有線性相關的最小平方法和加權之最小平方法來估計參數，而使用加權的目的，是為了使得參數估計時其估計量的變異數會因為不同的權數而呈現穩定。我們利用以上的方法將結果與 Bain 和 Engelhardt(1991)提出之最大概似估計量來進行比較。其步驟有三：

(a) 利用 Weibull 分配之累積分配函數為

$$F(x) = 1-\exp[-(\frac{x}{\eta})^{\beta}] \qquad (1.1)$$

經過二次對數轉換之後形成

$$\ln\{-\ln[1-F(x)]\} = \beta\ln x - \eta_0 \qquad (1.2)$$

的線性關係，其中 $\beta\ln\eta = \eta_0$。

而 Extreme-value 分配之累積分配函數為

$$F(x) = 1-\exp[-\exp(\frac{x-\xi}{\theta})] \qquad (1.3)$$

經過二次對數轉換之後形成

$$\ln\{-\ln[1-F(x)]\} = x\theta^{-1} - \xi_0 \qquad (1.4)$$

的線性關係，其中 $\xi_0 = \frac{\xi}{\theta}$。

(b) 在完整資料下，假設有 n 個隨機樣本 $X_1, X_2, \ldots, X_n$ 來自 Weibull 或是 Extreme-value 分配，而 $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ 為其對應之順序統計量，我們利用三種不同估計量 p_{1i}, p_{2i}, p_{3i} 來估計(1.2)及(1.4)式之 $F(x_{(i)})$，其中使用到 $p_{1i} = \frac{i}{n+1}$，$p_{2i} = \frac{i-0.5}{n}$，$p_{3i} = \frac{i-0.3}{n+0.4}$，$i = 1,\ldots,n$(參考 Kapur 和 Lamberson(1977)及 Ross (1987))。另外，考慮假設 $1-F(X)$ 是來自於均勻分配 $U(0,1)$，則 $-\ln[1-F(X)]$ 可視為指數分配 $Exp(1)$，因此利用 $p_{4i} = \sum_{j=1}^{i}\frac{1}{n-j+1}$，$i = 1,2,\ldots,n$ 估計(1.2)及(1.4)式中 $-\ln[1-F(x_{(i)})]$(參考 Lawless(1982)及 Ross(1978))。

而加權之權數 W_i 分別為

$W_{1ki} = [(1-p_{ki})\ln(1-p_{ki})]^2$、

$W_{2ki} = \frac{3.3p_{ki} - 275[1-(1-p_{ki})^{0.025}]}{n}$ 及

$W_i = (\sum_{j=1}^{i}\frac{1}{n-j+1})^2 = W_{3ki}, k = 1,2,3,$

$i = 1,\ldots,n$。

(c) 在第一失敗-設限抽樣方案下，有 m 個隨機樣本 $X_{(1)1}, X_{(1)2}, \ldots, X_{(1)m}$ 來自 Weibull 或是 Extreme-value 分配，而 $X'_{(1)1} < X'_{(1)2} < \ldots < X'_{(1)m}$ 為其對應之順序統計量，則(1.2)式可改寫為

$$\ln\{-\ln[1-F(x)]\} = \ln n + \beta\ln x - \eta_0 \qquad (1.5)$$

其中 $\beta\ln\eta = \eta_0$，而(1.4)式可改寫為

$$\ln\{-\ln[1-F(x)]\} = \ln n + x\theta^{-1} - \xi_0 \qquad (1.6)$$

其中 $\xi_0 = \frac{\xi}{\theta}$。我們利用 p_{1i}, p_{2i}, p_{3i} 來估計(1.5)及(1.6)式中 $F(x'_{(1)i})$，其中使用 $p_{1i} = \frac{i}{m+1}, p_{2i} = \frac{i-0.5}{m}, p_{3i} = \frac{i-0.3}{m+0.4}$，$i = 1,2,\ldots,m$。另外，若假設 $1-F(X)$ 來自於均勻分配 $U(0,1)$，則 $-\ln[1-F(X)]$ 可視為指數分配 $Exp(1)$，因此利用 $p_{4i} = \sum_{j=1}^{i}\frac{1}{m-j+1}$，$i = 1,2,\ldots,m$ 估計(1.2)及(1.4)式中之 $-\ln[1-F(x'_{(1)i})]$。

而加權之權數 W_i 分別為

$W_{1ki} = [(1-p_{ki})\ln(1-p_{ki})]^2$、

$W_{2ki} = \frac{3.3p_{ki} - 275[1-(1-p_{ki})^{0.025}]}{m}$ 及

$W_i = (\sum_{j=1}^{i}\frac{1}{m-j+1})^2 = W_{3ki}$，$k = 1,2,3$，$i = 1,2,\ldots,m$。

由上可知，我們將有 12 種點估計之方法來進行估計，並且將之整理成下表：

表 1　估計方法之總結

方法	形式	P	W
1	線性迴歸	p_{1i}	--
2	線性迴歸	p_{2i}	--
3	線性迴歸	p_{3i}	--
4	線性迴歸	p_{4i}	--
5	加權線性迴歸	p_{1i}	W_{11}
6	加權線性迴歸	p_{2i}	W_{12i}
7	加權線性迴歸	p_{3i}	W_{13i}
8	加權線性迴歸	p_{1i}	W_{21}
9	加權線性迴歸	p_{2i}	W_{22i}
10	加權線性迴歸	p_{3i}	W_{23i}
11	加權線性迴歸	p_{4i}	W_{34i}
12	最大概似法	--	--

最後，我們使用蒙地卡羅(Monte Carlo)模擬來評估前述之區間估計和點估計部分。此外，由於 Weibull 分配樣本取 ln 後可轉換為 Extreme-value 分配樣本，所以我們也將在模擬的部分進行討論。

三、結果與討論

整體來說，我們使用第一失敗-設限抽樣方案所模擬出的參數區間估計，除了顯著水準 $\hat{\alpha}$ 的平均數幾乎都很接近 α 的真值，表示準確度很高之外，標準差的估計值在兩個分配下都非常的小，表示我們估計的精確度也不錯。在 Weibull 分配方面，雙尾區間估計比單尾區間估計要來的準確；而 Extreme-value 分配方面，當樣本組數(m)增加，所估計顯著水準 $\hat{\alpha}$ 之標準差會隨之減少，表示精確度會趨於穩定。

另外，在參數的點估計方面，不論是完整資料或是第一失敗-設限抽樣方案 在大部分的情況下，我們所提出的最小平方估計法都會因為樣本數增加而提高估計的精確度和準確度。尤其是我們提出的最小平方加權點估計方法，幾乎都有不錯的效果，而又以最小平方加權點估計法中的方法 9 和方法 10，在兩種分配的任何情況下的表現幾乎都是非常不錯。

在模擬方法方面，Weibull 分配的點估計中，因為 Weibull 分配樣本的分散程度較大，所以我們建議採用將 Weibull 分配樣本取 ln 轉換為 Extreme-value 分配樣本之點估計方法進行模擬。而 Extreme-value 分配之區間估計中，因為參數 θ, ξ 的聯合信賴區間分別需要三條和四條的聯立方程式求解，在進行運算時容易發生溢位，所以我們建議採用將 Extreme-value 分配樣本取 exp 轉換為 Weibull 分配樣本來進行區間估計。

四、計劃成果自評

我們使用第一失敗-設限抽樣方案，對 Weibull 分配和 Extreme-value 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] Bain,L.J. and Engelhardt,M.(1992) , "*Introduction to Probability and Mathematical Statistics*", Duxbury Press.

[2] Balakrishnan,N. and Cohen,A.C. (1991), " *Order Statistics and Inference*", Academic Press.

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

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

[5] Billman,B.R., Antle,C.E and Bain,L.J.

(1972)"Statistic inference from cer sored Weibull samples", *Technometrics*, **14**, 831-840.

[6] Casella,G. and Berger,R.L.(1990), "*Statistical Inference*", Duxbury Press.

[7] Chen,Z.(1998),"Joint estimation for the parameters of Weibull distr bution", *Journal of Statistical Planning and Inference*, **66**, 113-120.

[8] Chen,Z.(1998),"Joint estimation for the parameters of the extreme value distributions", *Statistical Papers*, **39**, 135-146.

[9] Chao,A. and Hwang,S.J.(1986), "Comparison of confidence intervals for the parameters of Weibull and Extreme value distributions", *IEEE Transactions on Reliability*, **35**, 111-113.

[10] Cohen,A.C. and Whitten,B.J.(1988), "*Parameter Estimation in Relia-bility and Life Span Models*", New York: Marcel Dekker.

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

[12] Gumbel,E. J.(1958),"*Statistics of Extremes*", New York: Columbia University Press.

[13] Johnson,L.N. and Kotz,S.(1970), "*Distribution in Statistics : Continuous Univariate Distributions-Vol.1&2*", Wiley ,New York.

[14] Kapur,K.C. and Lamberson,L.R.(1977), "*Reliability in engineering design*",Wiley, New York.

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

[16] Lawless,J.F.(1982),"*Statistical Models and Methods for Lifetime Data*", John Wiley & Sons, New York.

[17] *Microsoft Developer Studio Fortran PowerStation 4.0 and IMSL*(1995) , Microsoft Corporation.

[18] Nordquist,J.M.(1945),"Theory of largest values, applied to earthquake magnitudes", *Transactions of American Geographical Union*, **26**, 29-31.

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

[20] Potter,W.D.(1949),"Normalcy tests of Precipitation and frequency studies of runoff on small watersheds", *U.S. Department of Agriculture Technical Bulletin*, No.**985**, Washington, DC: GPO.

[21] Rantz,S.F. and Riggs,H.C.(1949), "Magnitude and frequency of floods in the Columbia River Basin", *U.S. Geological Survey, Water Supply Paper* **1080**, 317-476.

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

[23] Ross,S.M.(1997),"*Simulation(2nd ed)*", Academic Press.

[24] Weibull,W.(1939),"A statistical theory of the material", Report No.**151** *Ingeniors Vetenskaps Akademiens Handligar*, Stockholm.

[25] Wu,J.W. and Tsai,W.L.(2000),"Failure Censored Sampling Plan for the Weibull Distribution", *International Journal of Information and Management Science*, **11(2)**, 13-25.