

可維修系統適合度之比較

呂理裕¹ 謝日章¹ 古美玉²

國立中央大學統計研究所¹

國立勤益技術學院 工業工程與管理系²

摘 要

本篇文章介紹一個可以用來描述可維修系統失效行為的模型 Power Law Process 以及這個模型的適合度檢定方法。將所收集到的資料做適當的轉換，再利用杜賓轉換消除資料中因為變異係數小所產生的超一致性後，比較 Lilliefors 檢定、Cramér-von Mises 檢定和一個較新發表的適合度檢定法 Total Time on Test Plot，簡稱 TTT Plot 三種方法的檢定力，想要探討新方法的檢定力是否高於其他兩種方法。最後利用蒙地卡羅法進行統計模擬研究。發現對所有的情形而言，TTT Plot 並不能說是一個比較好的檢定方法。但是在大部份的情形中，TTT Plot 檢定的表現明顯地要比 Lilliefors 檢定要好。所以如果要進行 Power Law Process 的適合度檢定，那麼 TTT Plot 應該是一個不錯的方法。

關鍵詞：可維修系統、杜賓轉換、超一致性

一、緒 論

在可維修系統(Repairable System)的壽命試驗上，要描述系統發生失效的情形，Ascher 和 Feingold (1984)推薦用 Power Law Process。在實務上，這是一個相當實用的模型。Duane(1964)把這個模型應用在通用電力公司的實作上得到不錯的結果。Power Law Process 不但可以用來描述平穩的系統(Stationary System)，更可以用來描述退化的系統(Deteriorating System)和改進的系統(Improving System)。這裡所謂一個退化的系統是指系統發生失效時相鄰的兩個失效間的時間間隔有越來越短的趨勢。反過來說一個改進的系統就是指系統相鄰的兩個失效間的時間間隔有越來越長的趨勢。而在一個平穩的系統中所觀察到的失效時間間隔則看不出有明顯增長或是減短的趨勢。

非齊性波氏過程(Nonhomogeneous Poisson Process)的重要特徵就是強度函數(Intensity Function)，是用來描述系統在每一個極短的時間間隔中會發生失效的機率。在符號上用 $\lambda(t)$ 來表示，其定義如下：

$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{\Pr[\text{在}(t, t + dt)\text{中有一個失效發生}]}{dt}$$

Power Law Process 其實是一個非齊性波氏過程的特例，較特別的是它的強度函數具有一種特定

的型式，如下所示：

$$f(t) = \frac{\beta}{\theta} \left(\frac{t}{\theta}\right)^{\beta-1}, \quad t > 0$$

這種型式是 Ascher(1981)所定義的，其中 β 和 θ 是兩個未知的參數。當 $\beta=1$ 的情況下，Power Law Process 就會被簡化成齊性波氏過程(Homogeneous Poisson Process)， θ 所代表的是相鄰失效間的平均時間。在這種情形下 Power Law Process 也就等於是一個更新過程(Renewal Process)，代表著系統處於平穩的狀態。當 $\beta > 1$ 時，強度函數的行為是隨著時間的流逝而遞增，那麼系統失效的機率也是隨著時間的流逝而逐漸遞增，也就是說此系統相鄰的兩個失效間的時間間隔會有越來越短的趨勢，所以稱此種系統是一個退化的系統。當 $\beta < 1$ 時，強度函數是隨著時間的流逝而遞減，則這個系統發生失效的機率也隨之遞減。同樣地可視為此系統相鄰的兩個失效間的時間間隔會有越來越長的趨勢，如此一來這個系統就是一個改進的系統。

在收集資料的方法上，常使用的方法有兩種。第一種方法稱之為失效截斷(Failure Truncated)。此種資料的收集方式為：當試驗進行至某一事先假設的失效個數發生時就停止收集資料。另外一種稱為時間截斷(Time Truncated)收集資料的方式為：當試驗進行到事前所預定的時間時即停止收集資料。本文做統計模擬研究時所生成的資料是採用失效截斷的方法。

在統計上，要對一組未知分佈的資料做統計推論前，通常都會先用適合度檢定(Goodness-of-Fit Test)找出資料所配適的分佈。Park 和 Kim(1992)所發表的論文中，提出比較 Kolmogorov-Smirnov 檢定、Cramér-von Mises 檢定和 Anderson-Darling 檢定的檢定力(Power)，結果是 Cramér-von Mises 方法的檢定力最高，本文則保留文獻中檢定力最高的方法再加上一個檢定指數分佈的 Lilliefors 檢定和一個較新發表的適合度檢定方法(Total Time on Test Plot)做比較，試圖探討這個新的方法是否較其它兩種檢定方法為優。至於評估一個檢定方法優劣的法則為：控制型 I 誤差的機率在 0.05 以下，比較以上三種適合度檢定方法的檢定力，看那一個檢定方法的檢定力較高即為較優之檢定方法。

二、適合度檢定方法

首先，假設 H_0 : 資料服從 Power Law Process，

H_1 : 資料不是服從 Power Law Process。

假設收集到一組系統失效的時間 T_1, T_2, \dots, T_n 。Moller(1976)建議將收集到的資料做對數比值轉換(Log-Ratio Transformation)後就可以做 Power Law Process 的適合度檢定。針對失效截斷型的資料其轉換形式如下：

$$w_i = -\ln \frac{T_{n-i}}{T_n}, \quad i=1, 2, \dots, n-1.$$

則轉換後的 w_i 's 是一組順序統計量 (Ordered Statistics) 服從平均數為 $\frac{1}{\lambda}$ 的指數分佈。

Rigdon(1989)說明經過上述的轉換後，只要是檢定指數分佈的適合度檢定方法都可以被引用來檢定 Power Law Process。

接下來依序介紹本文所採用的三種適合度檢定方法：

2.1 Lilliefors 檢定

Lilliefors(1969)提出一種指數分佈的適合度檢定方法。資料經過對數比值轉換之後就可以做 Power Law Process 的適合度檢定，先找估計的分佈 (estimated cumulative distribution function)。

$$F^*(w) = 1 - e^{-\frac{w}{\hat{\mu}_F}}, \quad w > 0.$$

$$\hat{\mu}_F = \sum_{i=1}^{n-1} \frac{w_i}{n-1}.$$

和經驗分佈 (empirical cumulative distribution function)

$$S(w) = \begin{cases} 0 & , w < w_1 \\ \frac{i}{n-1} & , w_i \leq w < w_{i+1} \quad , i=2, \dots, n-2. \\ 1 & , w \geq w_{n-1} \end{cases}$$

再找出估計分佈和經驗分佈之間最大的距離

$$D_1 = \max_{1 \leq i \leq n-1} |F^*(w_i) - S(w_i)|,$$

$$D_2 = \max_{1 \leq i \leq n-1} |F^*(w_i) - S(w_{i-1})|,$$

$$D = \max(D_1, D_2) \quad \circ$$

D 即是 Lilliefors 檢定的檢定統計量。再利用 Conover(1983)第 464 頁表中的臨界值，即可完成檢定。

2.2 Cramér-von Mises 檢定

Crow(1974)提出一個適合度檢定的方法，將收集到的資料做下列轉換： $z_i = \frac{T_i}{T_n}$ ，

$i=1, 2, \dots, n-1$ 。

經過轉換後的資料就可進行 Power Law Process 的適合度檢定，其步驟如下：

找出 的不偏估計式

$$\bar{\beta} = \frac{n-2}{n} \hat{\beta}$$

$$= \frac{n-2}{n} \frac{n}{\sum_{i=1}^{n-1} \ln \frac{t_n}{t_i}} .$$

($\hat{\beta}$ 的求法詳見 Rigdon 和 Basu(1989))。

計算修正的 Cramér-von Mises 統計量

$$C^2 = \frac{1}{2(n-1)} + \sum_{i=1}^{n-1} \left(z_i^{\bar{\beta}} - \frac{2i-1}{2(n-1)} \right)^2 ,$$

與 Crow(1974)所提供的臨界值做比較即可完成適合度檢定。

2.3 Total Time on Test Plot

這是由 B. Klefsjö 和 U. Kumar(1989)所提出的方法，假設收集到一組資料

$t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(n)}$ 要畫出這些觀察值的 Total Time on Test Plot 的步驟如下：

(1) 計算 TTT-values

$$S_i = nt_{(1)} + (n-1)(t_{(2)} - t_{(1)}) + \dots + (n-i+1)(t_{(i)} - t_{(i-1)}) ,$$

$i=1,2, \dots, n。$

假設 $S_0 = 0$ 。

(2) 將 TTT-values 標準化

$$y_i \equiv \frac{S_i}{S_n} \quad i=1,2, \dots, n。$$

(3) 描點 $\left(\frac{i}{n}, y_i\right)$, $i=1,2, \dots, n。$

(4) 將這些點連起來。

如果資料是來自 Power Law Process, 則 Total Time on Test Plot 畫出來的線應該會在單位正方形的對角線附近震盪, 如圖一所示。利用此種特性讓我們可以進行 Power Law Process 的適合度檢定。

定義：累積 TTT 統計量(Cumulative TTT Statistics)

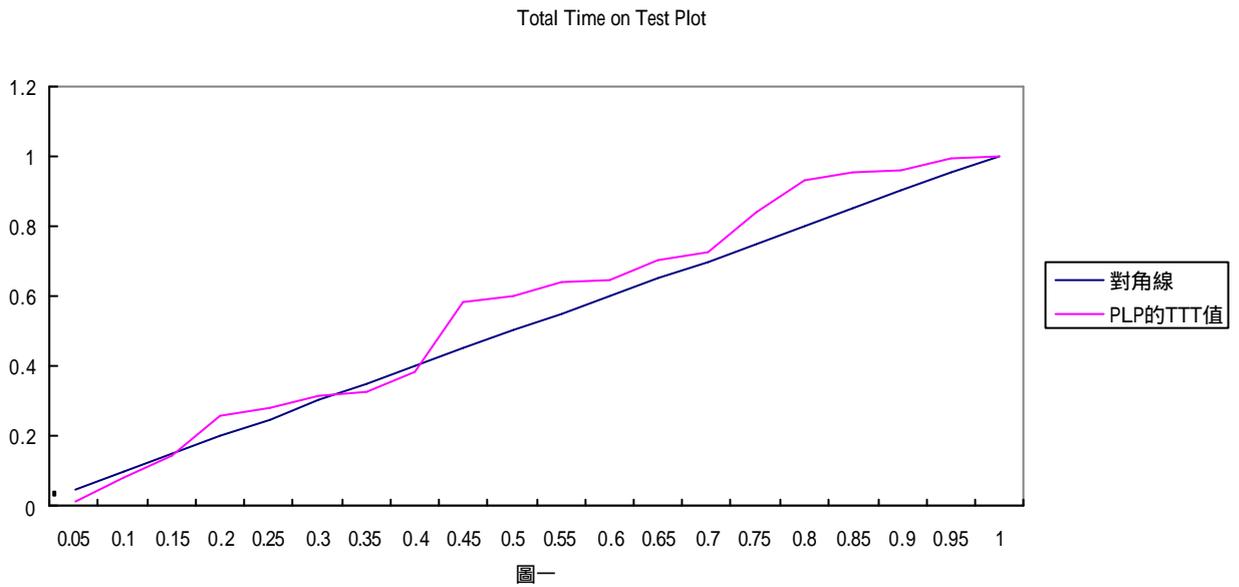
$$V \equiv \sum_{i=1}^{n-1} u_i .$$

其中 $u_i \equiv (n - i)(w_i - w_{i-1})$, $w_0 = 0$ 。

或者 $u_i \equiv (n - i)(d_i - d_{i-1})$, $d_0 = 0$ 。

其中 w_i 為經過對數比值轉換後的資料, d_i 是 w_i 再經過杜賓轉換 (Durbin Transformation) 後的資料(在下一章中介紹)。

V 統計量可以用來做 Power Law Process 的適合度檢定。如果畫出來的 TTT Plot 很接近對角



線, 則在 TTT Plot 以下所涵蓋的面積 A 也會接近 $\frac{1}{2}$ 。A 和 V 的線性關係為 $A = \frac{V + \frac{1}{2}}{n - 1}$,

所以 $A = \frac{1}{2}$ 時 $V = \frac{n}{2} - 1$ 。

當 n 夠大時, V 會近似 $N(\frac{n}{2} - 1, \frac{n - 2}{12})$,

所以 $\psi \equiv \frac{V - \frac{n}{2} + 1}{\sqrt{\frac{n - 2}{12}}} \sim N(0, 1)$ 。

以上三種方法皆是針對 Power Law Process 而發展的, 所以若需檢定資料是否服從 Power Law Process, 任選其中一種即可。本次研究試圖探討那一種方法的檢定力較高。

三、杜賓轉換

如果收集到的資料是來自一個更新過程，服從某一分佈。若此分佈的變異係數(Coefficient of Variation) $\frac{\sigma}{\mu}$ 小，則這些失效時間間隔的長度差不多都是相等的。這樣的結果會使得檢定統計

量因為資料的“超一致性(Superuniformity)”而造成檢定力降低。為要避免這種疏失，解決的方法就是去增加這些檢定方法的敏感度，也就是增加這些檢定方法的檢定力。此時可藉由 Durbin(1961)所提出的轉換解決這個困擾，其轉換的型式如下：

$$d_i = (n - i)(w_i - w_{i-1}) \quad , \quad i=1,2, \dots, n-1 ,$$

其中 $w_0 = 0$ 。

D'agostino 和 Stephens(1986)稱這些 d_i 's 是正規化的間距(Normalized Spacings) 資料中所隱含的超一致性已被杜賓轉換消除了，所以達到了提升檢定力的目的。

另外，Cramér-von Mises 法所檢定的資料並沒有做杜賓轉換，是因為 Lewis(1965)提到杜賓轉換是將資料從平均數為 $\frac{1}{\lambda}$ 的指數分佈轉換成平均數為 $\frac{1}{(n-i)\lambda}$ 的指數分佈，而 Crow 建議做的轉換並不是將資料轉換成指數分佈，所以 Cramér-von Mises 檢定並不適合做杜賓轉換。

四、統計模擬研究

在這一章中探討的主題是前面三種適合度檢定方法的檢定力。所以在對立假設成立的條件下，假設對立假設為一更新過程。我們找了三種在可靠度領域中常用壽命的分佈，以方便比較三種方法的檢定力。

1. Weibull 分佈。
2. Gamma 分佈。
3. Lognormal 分佈。

在 Power Law Process 的強度函數中參數 θ 代表著系統所處的狀態，所以考慮 $\theta = 1.5(>1)$ 、 $\theta = 1$ 和 $\theta = 0.5(<1)$ 三種情形。另外，第三章中曾提到檢定方法的檢定力與各分佈的變異係數有密切的關係，所以在做模擬時特別把變異係數分成三種情形來討論，分別如下：

1. 變異係數 = 0.5 (相當 $< \mu$)。
2. 變異係數 = 1 (相當 $= \mu$)。
3. 變異係數 = 1.5 (相當 $> \mu$)。

這樣一來，就可以清楚地看出系統狀態的改變和變異係數的變化分別對三種適合度檢定方法的影響。

在 Weibull 分佈以及 Gamma 分佈中，變異係數只和分佈的形狀參數(Shape Parameter)有關而和尺度參數(Scale Parameter)無關。而在 Lognormal 分佈中(利用常態分佈取指數為 Lognormal)，變異係數只和分佈中的尺度參數有關而和位置參數(Location Parameter)無關。所以，可求出對應於變異係數值的各個參數值，如下所示：

1. Weibull($v, 100$) v 為形狀參數。
 $v = 2.1$ 時 變異係數 = 0.5。
 $v = 1$ 時 變異係數 = 1。
 $v = 0.7$ 時 變異係數 = 1.5。
2. Gamma($\lambda, 100$) λ 為形狀參數。
 $\lambda = 4$ 時 變異係數 = 0.5。
 $\lambda = 1$ 時 變異係數 = 1。
 $\lambda = 0.44$ 時 變異係數 = 1.5。
3. Normal($3, \sigma$) σ 為尺度參數。
 $\sigma = 0.47$ 時 變異係數 = 0.5。
 $\sigma = 0.83$ 時 變異係數 = 1。
 $\sigma = 1.09$ 時 變異係數 = 1.5。

表一至表九是依照失效時間的分佈和資料的變異係數來分類並且列出各種檢定方法和樣本數不同時，三種適合度檢定方法所能達到的檢定力。

表一、Weibull 分佈 變異係數=0.5

方法 檢定力 樣本數	C. V. M. 檢定	TTT Plot 檢定		Lilliefors 檢定	
		無杜賓轉換	有杜賓轉換	無杜賓轉換	有杜賓轉換
10	0	0	0.569	0	0
20	0	0	0.958	0	0
30	0	0	0.997	0	0
40	0	0	1	0	0
50	0	0	1	0	0
60	0	0	1	0	0

表二、Weibull 分佈 變異係數=1

方法 檢定力 樣本數	C. V. M. 檢定	TTT Plot 檢定		Lilliefors 檢定	
		無杜賓轉換	有杜賓轉換	無杜賓轉換	有杜賓轉換
10	0.059	0.026	0.043	0	0
20	0.056	0.033	0.061	0	0
30	0.055	0.034	0.049	0	0.005
40	0.053	0.036	0.040	0.008	0.020
50	0.054	0.039	0.050	0.025	0.051
60	0.054	0.041	0.057	0.054	0.099

表三、Weibull 分佈 變異係數=1.5

方法 檢定力 樣本數	C. V. M. 檢定	TTT Plot 檢定		Lilliefors 檢定	
		無杜賓轉換	有杜賓轉換	無杜賓轉換	有杜賓轉換
10	0.301	0.102	0.215	0	0
20	0.292	0.112	0.448	0.009	0.021
30	0.290	0.129	0.598	0.053	0.092
40	0.286	0.139	0.749	0.131	0.204
50	0.288	0.137	0.819	0.230	0.313
60	0.288	0.158	0.875	0.328	0.423

表四、Gamma 分佈 變異係數=0.5

方法 檢定力 樣本數	C. V. M. 檢定	TTT Plot 檢定		Lilliefors 檢定	
		無杜賓轉換	有杜賓轉換	無杜賓轉換	有杜賓轉換
10	0	0	0.646	0	0.001
20	0	0	0.981	0	0
30	0	0	0.999	0	0
40	0	0	1	0	0
50	0	0.001	1	0	0
60	0	0.001	1	0	0

表五、Gamma 分佈 變異係數=1

方法 檢定力 樣本數	C. V. M. 檢定	TTT Plot 檢定		Lilliefors 檢定	
		無杜賓轉換	有杜賓轉換	無杜賓轉換	有杜賓轉換
10	0.047	0.025	0.041	0.075	0.142
20	0.050	0.038	0.054	0.058	0.126
30	0.050	0.038	0.062	0.054	0.113
40	0.048	0.038	0.050	0.053	0.108
50	0.049	0.043	0.054	0.054	0.096
60	0.054	0.048	0.056	0.054	0.099

表六、Gamma 分佈 變異係數=1.5

方法 檢定力 樣本數	C. V. M. 檢定	TTT Plot 檢定		Lilliefors 檢定	
		無杜賓轉換	有杜賓轉換	無杜賓轉換	有杜賓轉換
10	0.272	0.153	0.361	0.324	0.417
20	0.325	0.160	0.633	0.367	0.490
30	0.337	0.168	0.815	0.381	0.501
40	0.334	0.181	0.903	0.381	0.505
50	0.345	0.169	0.960	0.396	0.502
60	0.348	0.186	0.985	0.396	0.499

表七、Lognormal 分佈 變異係數=0.5

方法 檢定力 樣本數	C. V. M. 檢定	TTT Plot 檢定		Lilliefors 檢定	
		無杜賓轉換	有杜賓轉換	無杜賓轉換	有杜賓轉換
10	0	0	0.766	0	0
20	0	0	0.983	0	0
30	0	0	0.998	0	0
40	0	0.001	1	0	0
50	0	0.001	1	0	0
60	0	0.001	1	0	0

表八、Lognormal 分佈 變異係數=1

方法 檢定力 樣本數	C. V. M. 檢定	TTT Plot 檢定		Lilliefors 檢定	
		無杜賓轉換	有杜賓轉換	無杜賓轉換	有杜賓轉換
10	0.024	0.014	0.100	0.041	0.077
20	0.031	0.013	0.212	0.039	0.075
30	0.036	0.025	0.262	0.040	0.068
40	0.037	0.032	0.302	0.041	0.074
50	0.037	0.034	0.377	0.043	0.065
60	0.038	0.037	0.403	0.039	0.063

表九、Lognormal 分佈 變異係數=1.5

方法 檢定力 樣本數	C. V. M. 檢定	TTT Plot 檢定		Lilliefors 檢定	
		無杜賓轉換	有杜賓轉換	無杜賓轉換	有杜賓轉換
10	0.100	0.066	0.073	0.140	0.206
20	0.144	0.078	0.151	0.171	0.245
30	0.170	0.112	0.200	0.201	0.273
40	0.191	0.116	0.243	0.220	0.285
50	0.199	0.125	0.277	0.230	0.297
60	0.218	0.136	0.329	0.249	0.312

五、 結 論

因為 Cramér-von Mises 檢定並沒有進行杜賓轉換，所以這個方法無法診斷並消除資料中所隱含的超一致性，所以在變異係數小的時候，Cramér-von Mises 檢定的檢定力都偏低 Cramér-von Mises 檢定這明顯是被資料中隱含的超一致性所影響。不過，亦可以在前面的表中發現，在不同的壽命分佈中 Cramér-von Mises 檢定的檢定力有隨著變異係數的增大而漸漸升高的情形。這個結果恰巧可以說明了此種檢定方法無法偵測出資料中所隱含的超一致性，所以當超一致性隨著變異係數的增大而漸漸減弱時，Cramér-von Mises 檢定的檢定力也就漸漸地提升了。所以，在超一致性效應很明顯時，欲檢定資料是否服從 Power Law Process，並不建議使用 Cramér-von Mises 檢定。

接下來就直接比較 TTT Plot 檢定和 Lilliefors 檢定。在第四章的表一至表九裡可以看出經過杜賓轉換之後兩種檢定方法的檢定力至少不會比轉換前的檢定力低甚至還會比轉換前高出許多。這是因為將資料中所隱含的超一致性去除後，檢定方法更容易判別所得到的資料是否來自於 Power Law Process。如果再把變異係數所造成的影響列入考慮的話，可在表中看出當變異係數等於 0.5 時 TTT Plot 檢定在不同的分佈中都算是較好的檢定方法。當變異係數等於 1 的時候，TTT Plot 檢定的表現只有在 Gamma 分佈時比 Lilliefors 檢定要差。當變異係數為 1.5 時，TTT Plot 的檢定力只有在壽命分佈是 Lognormal 時略遜於 Lilliefors 檢定。但是隨著樣本數的增大，TTT Plot 檢定的表現會有比 Lilliefors 檢定要好的情形發生。由此可知，TTT Plot 檢定在變異係數小時表現較佳，即說明了 TTT Plot 檢定比較不受超一致性的影響，算是比較穩健的方法。所以，本次研究建議使用 TTT Plot 檢定來進行 Power Law Process 的適合度檢定。

總之，對所有的情況來說 TTT Plot 檢定並不是最好的檢定方法。但是在大部份的情形中，TTT Plot 檢定的表現明顯地要比 Lilliefors 檢定要好。所以如果要做 Power Law Process 的適合度檢定，那麼 TTT Plot 應該是一個值得考慮的方法。

參 考 文 獻

- [1] H. Ascher, H. Feingold, "Repairable Systems Reliability Modeling, Inference, Misconceptions and Their Causes", Marcel Dekker, 1984
- [2] W. J. Conover, "Practical Nonparametric Statistics", 華泰書局, 1983.
- [3] L. H. Crow, "Reliability Analysis for Complex Repairable Systems", Reliability and Biometry, 1974, pp.379-410.
- [4] R. B. D'Agostino, M. A. Stephens, "Goodness-of-Fit Techniques", Marcel Dekker, 1986.
- [5] J. T. Duane, "Learning Curve Approach to Reliability", IEEE Transaction on Aerospace 2, 1964, pp.563-566.
- [6] J. Durbin, "Some Methods of Constructing Exact Tests", Biometrika48, 1 and 2, 1961, pp.41-55.
- [7] B. Klefsjö, U. Kumar, "Goodness-of-Fit for the Power Law Process Based on the TTT Plot", IEEE Transactions on Reliability, Vol.41, No.4, 1989, pp.251-260.
- [8] P. A. W. Lewis, "Some Results on Tests for Poisson Processes", Biometrika, 52, 1 and 2, 1965, pp.67-77.
- [9] H. W. Lilliefors, "On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown", Journal of the American Statistical Association 64, 1969, pp.387-390.
- [10] S. K. Moller, "The Rasch-Weibull Process", Scandinavian Journal of Statistics, 3, 1976, pp.107-115.
- [11] W. J. Park, Y. G. Kim, "Goodness-of-Fit Tests for the Power Law Process", IEEE Transactions on Reliability, Vol.41, No.1, 1992, pp.107-111.
- [12] S. E. Rigdon, A. P. Basu, "The Power Law Process: A Model for the Reliability of Repairable System", Journal of Quality Technology, Vol.21, No.4, 1989, pp.251-260.
- [13] S. E. Rigdon, "Testing Goodness-of-Fit for the Power Law Process", Commun. Statist.-Theory Meth. 18(12), 1989, pp.4665-4676.
- [14] S. M. Ross, "A Course In Simulation", Macmillan Publishing Company, 1990.