Computer Graphic에 의한 와이블분포 모수추정에 관한 연구
- A study on the Parameter Estimation of the Weibull Distribution using Computer Graphic Method -

嚴 泰元 *
鄭 秀一 **

ABSTRACT

This study deals with the estimation of the Weibull parameters, which have a close relation with product reliability characteristics.

Among the many kinds of estimation methods, Kao's Weibull Probability Paper(WPP) is commonly used. The WPP is very convenient, but it has a great disadvantage in estimation accuracy by plotting method. It is very difficult to get the same results even if one use the same data several times.

A computer program for the regression method is used for the parameter estimation to reduce these errors.

Especially, the computer graphic program was written in GW-BASIC 3.22 language and the program appears in the appendix part with a couple of running examples for user's reference.

I. 서 론

본 연구는 최근 품질보증및 제품책임에 대한 관심이 높아짐에 따라 많은 제품의 신뢰성 특성을 나타내 주는 와이블분포의 모수들을 보다 정확하고 신속히 추정함으로써 그들 제품의 신뢰성관리를 보다 효율적으로 이끌어 나가는데 도움을 주고자 하였다.

분포의 모수를 추정하는 데는 여러가지 방법이 있으나 이들 중 가장 손쉽고 편리한 와이블확률지를 택하였으며 여러종류의 확률지중에서도 α, β, γ scale이 표시되어 있어 모수추정을 가장 용이하게 할 수 있는 石川의 형식을 참고 하였다.

그리고 와이블확률지의 맹점으로 지적되고 있는 plotting한 점들을 직선 또는 곡선으로 fitting시키는데 있어서의 개개인 및 작도상의 오차를 최대한으로 줄이기 위한 line fitting을 컴퓨터 프로그램으로 처리하였다. 또한 와이블확률지에 의한 모수추정치와 컴퓨터 프로그램을 이용한 모수추정치의 효율을 비교, 검토하였다.

II. 이론적 고찰

1. 와이블분포 [2], [4]

와이블분포는 많은 실제적인 현상으로부터 일어나는 time-to-failure데이타를 fitting시키는 데 있어

* 柳韓專門大學 工業經營科 副敎授
** 仁荷大學校 產業工學科 敎授
접수 : 1993년 4월 29일
확정 : 1993년 5월 10일

서의 다양한 융통성으로 인하여 신뢰성, 수명시험 등의 분야에 넓게 이용되고 있으며 또한 Weibull plot 가 이질의 또는 mixed distribution들을 보여 주는데 민감하기 때문에 지수분포, 정규분포들도와이블 분포에 의해 근사시킬 수가 있다.
와이블분포는 다음과 같은 누적분포함수 F(t)및 확률밀도함수 f(t)를 갖는 분포이다.

$$F(t)=1-R(t)=1-\exp\left[\frac{-(t-\gamma)^{\beta}}{\alpha}\right]$$

$$f(t)=\frac{\beta(t-\gamma)^{\beta-1}}{\alpha}\cdot\exp\left[\frac{-(t-\gamma)^{\beta}}{\alpha}\right]$$

윗 식에는 모수가 3개 있고, 각 각 다음과 같이 정의된다.
α : 척도의 모수 (scale parameter)
β : 형의 모수 (shape parameter)
γ : 위치의 모수 (location parameter)

$y=\frac{t-\gamma}{\alpha^{\frac{1}{\beta}}}$ 로 놓고 y의 확률밀도함수 f(y)를 구하여 보면 $dt=\alpha^{\frac{1}{\beta}}dy$가 되므로

$$f(t)dt=\beta\left[\frac{t-\gamma}{\alpha^{\frac{1}{\beta}}}\right]\cdot\exp\left[-\frac{t-\gamma}{\alpha^{\frac{1}{\beta}}}\right]\cdot\frac{dt}{\alpha^{\frac{1}{\beta}}}$$

$$=\beta y^{\beta-1}\cdot\exp[-y^{\beta}]\,dy$$

$$\therefore f(y)=\beta y^{\beta-1}\cdot\exp[-y^{\beta}]$$

이상으로 γ나 α는 어느것이나 시간축의 대소에만 관계되는 모수로서 본질적인 모수는 β라는 것을 알 수 있다. 즉 β가 크면 확률밀도함수가 뾰족해짐과 동시에 모든 제품의 특성이 비슷하다는 것을 나타낸다. 또한 $\beta>1$, $\beta=1$, $\beta<1$에 따라서 고장의 발생상태가 고장률증가형, 고장률일정형, 고장률감소형과 같이 형상적으로 구별되는 것도 와이블분포의 편리한 점이라 할 수 있다.

그런데 데이터를 와이블분포에 맞추어서 취급하려면 α, β, γ의 세가지 모수를 추정하지 않으면 않된다. 이들이 추정되면 β의 값에 따라 고장의 성질이 파악 된다거나, 임의의 시간 t까지의 신뢰도 등을 알게 된다.

실측데이터로 부터 이들의 모수를 추정하는 가장 실용적인 방법은 와이블확률지를 사용하는 것이다.

2. 와이블확률지의 구성 및 모수추정방법 [2], [4]

와이블확률지의 구성은 다음과 같이 되어 있다.

$$F(t)=1-\exp\left[\frac{-(t-\gamma)^{\beta}}{\alpha}\right]$$

윗식에서 $\gamma=0$으로 놓고 이항하면

$$1-F(t)=\exp\left[\frac{-t^{\beta}}{\alpha}\right]$$

가 된다. 이것을 $\ln\frac{1}{1-F(t)}=\frac{t^{\beta}}{\alpha}$로 변형하고 다시 한번 양변의 자연대수를 취하면

$$\ln\ln\frac{1}{1-F(t)}=\beta\ln t-\ln\alpha$$

가 된다. 변수 t를 포함한 항을

$$\ln\ln\frac{1}{1-F(t)}=Y,\quad \ln t=X$$

로 생각하면 윗 식은 Y = AX + B 와 같은 형이된다.

확률지는 $\ln\ln\frac{1}{1-F(t)}$ 가 종축에 $\ln t$ 가 횡축에 등간격 눈금으로 매겨져 있다. $\gamma=0$인 와이블분포를 따르는 데이터를 플로트하면 그 경사가 형의 모수 β, 절편이 $-\ln\alpha$ 인 직선이 된다.
어떤 시간 t(확률지의 아래 눈금)에서 n개중 r번째의 고장이 났다면 t와 F(t) 퍼센트(확률지의 좌측)의 눈금으로 F(t)=r/(n+1)에 대응하는 점에 타점한다. 이렇게 하여 구한 점들에 가장 알맞는 직선을 긋고, 경사를 종축(확률지의 우측)의 눈금을 사용하여 읽으면 이것이 β이다. 또 α는 데이터에 적합시킨 직선이 t=1의 종의 주축과 만나는 점으로 부터 우측으로 횡축과 평행선을 그어 우측의 lnln 1/[1-F(t)]의 눈금을 읽으면 이것이 $-\ln\alpha$ 이다. $-\ln\alpha$를 α로 고치기 위해서는 횡축의 위의 눈금과 아래의 눈금이 ln t와 t와의 관계를 나타내고 있으므로 ln α의 값을 위 눈금에서 취하고, 그 점으로 부터 똑바로 내려와 t를 읽으면 이것이 α가 된다.

III. 컴퓨터 프로그램의 작성

1. 프로그램의 개요

본 연구에서 제시된 와이블분포의 모수를 추정하는 프로그램의 내용은 대별해서 γ가 0일 경우, γ의 값이 주어졌을 경우, γ의 값까지도 추정해야 하는 경우로 나눌 수 있으며 추정기법으로서 최소자승법을 사용하였다.[1] 즉, 추정 회귀식의 일반식을

$$Y_i=b_0+b_1x_1+\cdot\cdot\cdot\cdot\cdot+b_n\,x_i^n$$

이라 하면 오차자승의 합은

$$S=\sum_{i=1}^{n}(Y_i-y_i)^2=\sum_{i=1}^{n}(b_0+b_1x_1+\cdot\cdot\cdot\cdot\cdot+b_n\,x_i^n-y_i)^2$$

이 되고 이를 최소로 하기 위한 방정식의 해는 다음 식으로 부터 구할 수 있다.

$$\frac{\partial S}{\partial b_0}=2\sum_{i=1}^{n}(b_0+b_1x_1+\cdot\cdot\cdot\cdot\cdot+b_n\,x_i^n-y_i)=0$$

$$\frac{\partial S}{\partial b_1}=2\sum_{i=1}^{n}(b_0+b_1x_1+\cdot\cdot\cdot\cdot\cdot+b_n\,x_i^n-y_i)\,x_1=0$$

$$\vdots$$

$$\frac{\partial S}{\partial b_n}=2\sum_{i=1}^{n}(b_0+b_1x_1+\cdot\cdot\cdot\cdot\cdot+b_n\,x_i^n-y_i)\,x_i^n=0$$

그러나 와이블확률지상에서 γ=0인 경우에는 1차식까지, γ>0인 경우에는 2차식까지만 이용해도 충분하므로 위의 방적식을 2차식까지만 풀어서 Y_i의 계수 $b_0,\,b_1,\,b_2$를 구하였으며 이를 이용 다음과 같이 프로그램을 구성하였다.

1) 필요한 데이터를 읽어 들이는 부분
2) γ가 0일 때 α, β를 추정하는 부분
3) γ를 주었을 때 α, β를 추정하는 부분
4) γ를 모를 때 α, β, γ를 추정하는 부분
5) Problem Identification을 인쇄해 주는 부분
6) 결과치를 인쇄해 주는 부분
7) 모수추정의 결과를 PC 화면상에서 그래픽으로 보여주는 부분

γ를 추정할 경우에는 error(각 점들의 추정회귀곡선으로 부터의 편차 제곱의 합)가 가장 작은 두 γ값 사이에 정확한 γ값이 존재한다는 가정하에 시간 t=0에서 부터 처음으로 고장이 나는 시간 OX(1)까지를 잘게 나누어 각 각의 구간마다 γ값을 지정해 주고 regression을 이용해 error가 가장 작은 경우의 γ값을 추출한 후 이 γ값으로 부터 OX(1)까지를 또다시 세분해 유효숫자 자리 만큼 반복계산을 하도록 하였다.

2. 데이터의 입력방법

데이터의 입력은 추정하려는 문제의 성격상 A) $\gamma=0$ 인 경우 B) γ값이 주어질 경우 C) γ값을 추정해야 할 경우로 나눌 수 있으며 기본적으로는 화면상의 지시 내용대로 입력시켜 나가면 된다.

다음은 프로그램을 실행시켰을 경우의 화면의 구성내용을 알아본 것이다.

1) 결과치를 화면으로 볼것인지 아니면 프린터로 인쇄할 것인지를 물어본다
   화면인 경우 : 1　　　　　　　　프린터인 경우 : 2
2) 지금 사용하는 PC의 모니터가 color인지 mono인지를 물어본다. 이는 graphic을 할 때의 화면의 해상도 차이를 인식해 화면을 구성하기 때문이다. 기본적인 video card의 해상도는 640 x 350 이상이어야 한다
   color인 경우 : 9를 입력　　　　　mono인 경우 : 3을 입력
3) 입력자료를 data file로 미리 작성하여 프로그램을 실행시킬 것인지 아니면 직접 key board를 이용해 데이터를 입력할 것인지의 선택
   key board 인 경우 : K　　　　　　data file 인 경우 : D
4) γ값의 추정여부
   -값 : γ가 주어질 경우
   0값 : γ가 0인 경우
   +값 : γ를 추정하는 경우
5) Input data 이외의 점에서의 신뢰도 R(t)의 계산
   +값 : 구한다
   그 밖의 경우에는 구하지 않는다
6) F(i)가 주어지는 지의 여부
   +값 : 주어진다
   그 밖의 경우에는 선택된 타점방법에 따라 컴퓨터가 계산
   · 평균　rank법으로 계산하는 경우 : 1
   · 메디안 rank법으로 계산하는 경우 : 2
   · 누적 hazard법으로 계산하는 경우 : 3

IV. 와이블확률지에 의한 결과와 컴퓨터 프로그램에 의한 결과의 비교

	예 1		예 2		예 3	
α	23.45		21.80		175,722.10	
β	1.00		1.90		1.62	
γ	0.00		2.00		1,249.00	
	t	F(t)	t	F(t)	t	F(t)
	2.76198	0.111	3.15732	0.066	1,400	0.02
	5.89326	0.222	5.85027	0.334	1,600	0.07
	9.50803	0.333	7.35716	0.714	1,800	0.13
	13.78340	0.444	9.32857	0.879	2,000	0.21
	19.01650	0.556	11.52348	0.969	2,400	0.44
	25.76260	0.666	13.57360	0.989	3,000	0.70
	35.27090	0.778	15.75914	0.999	4,000	0.88
	51.52500	0.889			5,000	0.96

표 1. 프로그램 성능검토용 고장시간 데이터

표에 제시된 고장시간 t에 대한 자료는 실측 데이터가 아니고, 프로그램의 성능검토 및 graphical method와의 비교를 위해 $t = \exp\left\{ \frac{\ln(-\ln R(t)) + \ln\alpha}{\beta} \right\} + \gamma$ 에 미리 지정된 α, β, γ와 R(t)를 대입하여 구한 값이다.

표1의 데이터를 이용하여 Graphical method와 Computer method에 의해 모수들을 추정한 결과 다음과 같았다.

	예 1		예 2		예 3	
	Graphical	Computer	Graphical	Computer	Graphical	Computer
α	24.0	23.45	25.20	21.78	169,460.595	175,722.056
β	1.0	1.0	2.02	1.89	1.610	1.622
γ	0.0	0.0	2.00	2.00	1,250.000	1,248.800

표 2. graphical method와 computer method의 결과치 비교

예1 과 예2의 경우 거의 무시할 수 있을 정도의 차이를 보여 주었다. 그러나 예 3의 γ까지도 추정해야 될 경우에는 비교적 큰 오차를 수반하였다. 그 원인을 분석해 본 결과, 역시 γ를 추정하는 작도과정에서의 오차가 주원인이었다.

V. 결 론

Graphical method에 의한 모수추정의 경우, β가 0일 때와 β의 값이 주어졌을 때에는 error가 비교적 작았으나 β의 값까지도 추정해야 할 경우에는 데이터의 자릿수가 많은데에도 원인이 있겠으나 α의 값이 너무나 작게 나왔다. 그러나 컴퓨터에 의한 결과는 정확하게 이론적인 수치를 얻을 수 있었다. 그러나 각 각의 추정방법에 있어서 데이터의 상태에 따라 그 효율이 달라진다. 예를 들면 least-square estimation은 transform data에서, maximum likely hood estimation은 group data에서 효율이 좋다고 알려져 있다.[8] 그러므로 차후에 데이터의 상태에 구애받지 않고도 효율이 우수한 방법이 연구되어야 하겠다.

참고문헌

[1] 박경수, 신뢰성공학 및 정비이론, 탑출판사, 1978

[2] 이근희, 완전품질관리, 창지사, 서울, 1981

[3] ______, 품질관리(이론과 실제), 상조사, 서울, 1990

[4] 황의철, 최신품질관리, 무역경영사, 1980

[5] Bain, L. J. and Antle, C.E., Estimation of Parameters in the Weibull Distribution, *Technometrics*, Vol 9, 1967

[6] Berrettoni, J.N. Practical Application of the Weibull Distribution, *Industrial Quality Control*, Vol 21, 1964

[7] Ireson, W.G.,*Reliability Handbook*, McGrawhill Book Company, 1966

[8] Kao, J.H.K., Computer Methods for Estimating Weibull Parameters in Reliability Studies, *Transactions of IRE-Reliabiliity and Quality Control*, Vol 13, 1958

[9] Kao, J.H.K, A Graphical Estimation of Mixed Weibull Parameters in Life Testing Electron Tubes, *Technometrics*, Vol 1, 1959

[10] Kimball, B. J., On the Choice of Plotting Positions on Probability Paper, *Journal of the American Statistical Association*, Vol 55, 1960

[11] Nelson, L. S., Weibull Probability Paper, *Industrial Quality Control*, Vol 23, 1967