

# 극치분포의 모수추정에 관한 연구

## A Study on the Parameter Estimation of Limiting Distribution

嚴泰元\*

### Abstract

This paper deals with the estimation of limiting distribution parameters, which have a close relation with product reliability characteristics in reliability test.

Among the many kinds of estimation methods, probability paper is commonly used. The limiting probability paper 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 a couple of running examples for user's reference appears in the appendix part.

### I. 서론

본 연구는 그동안 발표된 와이블 확률지와 누적 해저드지의 모수추정에 이은 확률지를 이용한 모수추정 프로그램 개발의 일환으로 신뢰성 테스트시 극치 확률지의 사용을 보다 효율적이고 간편하게 할 수 있도록 하기 위한 연구목적으로 수행되었다.

한편 본 연구에서는 확률지 사용시 맹점으로 지적되고 있는 타점한 점들을 직선 또는 곡선으로 적합시키는 데 있어서의 개개인 및 작도상의 오차를 최대한으로 줄이기 위한 선적합(line fitting)을 컴퓨터 프로그램으로 처리하였으며 극치 확률지에 의한 모수추정치와 컴퓨터 프로그램을 이용한 모수추정치의 효율을 비교, 검토하였다.

### II. 이론적 고찰

#### 1. 극치분포

극치분포에 관한 이론은 1928년 Fisher와 Tippett의 논문에서 제안되어 1958년 Gumbel에 의해 정리되었으며[8] 만약  $G(x)$ 가 극치분포에 따른다면 분포함수는

$$G^k(x) = G(a_k x + b_k) \quad \text{단, } a_k, b_k \text{는 상수}$$

(2.1)

로 정의되며[8] 대표본일 경우  $x$ 와  $a_k x + b_k$ 는 선형관계를 나타낸다.

---

\* 柳韓專門大學 工業經營科

표 2.1 프로그램 성능검토용 시험데이터

예 1 : 최대 분포	예 2 : 최소 분포	예 3 : 최소 분포
중격 : Kg	길이 : mm	시간 : hour
12	1213	116
5	1220	73
16	1224	46
9	1207	30
10	1218	19
13	1215	7
7		
20		
8.5		
15		

또한 극치분포의 형태는 Fisher와 Tippett에 의해 다음과 같은 형태로 제안되었다.[9]

$$F(x) = \exp\{-\exp(-y)\} \quad \text{단, } y = \alpha(x-u)$$

(2.2)

여기서  $\alpha$ 는 범위의 모수, 일명 Gumbel의 기울기라고도 하며 모수  $u$ 는 분포의 모드로서 극치분포의 평균과 분산은 다음 식으로 얻어진다.[9]

$$\mu = u + \gamma/\alpha \tag{2.3}$$

$$\sigma^2 = (\pi^2/6)\alpha^2 \quad \text{단, } \gamma = \text{Euler의 상수 } (\approx 0.5772 \dots) \tag{2.4}$$

## 2. 극치확률지

### 2.1 극치확률지의 구성

극치 확률지는 다음과 같이 구성되어 있다.[2]

식 (2.2)의 분포함수

$$F(x) = \exp\{-\exp(-y)\}$$

의 양변의 ln을 거듭 취하면

$$\begin{aligned} y &= -\ln \ln(1/F(x)) \\ &= \alpha(x-u) \end{aligned}$$

이 된다. 여기서  $x$ 축을 등간격으로 하고  $y$ 축을  $-\ln \ln(1/F(x))$ 인 눈금을 매긴 그래프가 바로 극치 확률지이다. 따라서 극치분포의 조건을 만족시키는 데이터는 이 확률지상에 타점했을 경우 직선의 형태를 따른다는 것을 나타낸다.

### 2.2 극치분포의 모수추정방법

극치분포의 모수추정방법은 크게 최대값에서의 추정과 최소값에서의 추정으로 나눌 수 있다. 두 추정방법간의 차이는 확률지의  $x$ 축상의 값을 반대로 해서 사용하는 차이밖에는 없다. 따라서 최대값에서의 추정방법을 먼저 알아보고 그 차이점을 기준으로 최소값에서의 추정방법을 알아보기로 한다.

식 (2.2)의

$$F(x) = \exp\{-\exp(-y)\} \quad \text{단, } y = \alpha(x-u)$$

로부터  $x=u$ 일 때  $y=0$ 이고  $F(u) = e^{-1}$ 이다. 이는 확률지상의  $y$ 축 눈금이 0일 때 직선과의 교점인  $x$ 축상의 눈금을 읽으면  $u$ 의 추정값이 되며, 식  $\mu = u + \gamma/\alpha$ 로부터  $\gamma = \alpha(\mu - u)$ 이고  $x=u$ 일 때  $y = \alpha(\mu - u) = \gamma$ 가 되므로  $y$ 축 눈금이  $\gamma$ 일 때 직선과의 교점인  $x$ 축상의 눈금을 읽으면  $\mu$ 의 추정값이 된다는 것을 의미한다.

표 2.2 그래픽방법과 컴퓨터방법의 결과치 비교

	예 1		예 2		예 3	
	그래픽방법	컴퓨터방법	그래픽방법	컴퓨터방법	그래픽방법	컴퓨터방법
$\alpha$	0.247	0.243	0.164	0.178	0.029	0.021
$u$	9.600	9.402	1219.500	1218.970	54.598	70.666
$\mu$	11.900	11.772	1216.000	1215.734	29.964	43.387
$\sigma^2$	26.960	27.743	61.160	51.696	1955.925	3673.868

또한 범위의 모수  $\alpha$ 는 직선상의 임의의 두점을  $(x_1, y_1), (x_2, y_2)$ 라 할 때  $y_1 = \alpha(x_1 - u), y_2 = \alpha(x_2 - u)$ 로 부터  $\alpha = (y_2 - y_1) / (x_2 - x_1)$ 이 된다.

한편 최소값에서의 모수추정방법은 최대값에서 사용하는 확률지를 그대로 사용하며  $x$ 축상의 눈금을 반대로 매겨서 사용하면 되므로  $\alpha$ 와  $\mu$ 의 추정식이 각각  $\alpha = (y_1 - y_2) / (x_1 - x_2)$ 로  $\mu = u - \gamma / \alpha$ 로 바뀌는 외에는 최대값에서의 추정방법과 같게 된다.

### III. 컴퓨터 프로그램의 작성

본 연구에서 제시된 극치분포의 모수를 추정하는 프로그램의 내용은 대별해서 최대분포 문제와 최소분포 문제로 나눌 수 있으며 추정기법으로서 최소자승법을 사용하였다.

즉, 추정 회귀식의 일반식을

$$Y_i = b_0 + b_1 x_1 + \dots + b_n x_i^n$$

(3.1)

이라 하면 오차자승의 합은

$$S = \sum_{i=1}^n (Y_i - y_i)^2 = \sum_{i=1}^n (b_0 + b_1 x_1 + \dots + b_n x_i^n - y_i)^2$$

(3.2)

이 되고 이를 최소로 하기 위한 방정식의 해를 풀어  $Y_i$ 의 계수  $b_0, b_1, b_2$ 를 구하는 프로그램을 작성하였다.

본 프로그램은 PC사용자들을 위해 작성하였기 때문에 초보자들도 쉽게 이용할 수 있도록 대화식 메뉴방식을 사용하였다.

즉, 입력데이터의 방법을 데이터 화일(data file)로 할 것인지 또는 컴퓨터 자판(keyboard)에서 직접 입력할 것인지의 여부, 결과치 인쇄를 컴퓨터 화면(monitor)으로 할 것인지 또는 프린터(printer)로 할 것인지, 고장비율(failure percentage) 계산을 어떤 방법으로 할 것인지 등을 메뉴에 따라 사용자가 선택적으로 사용할 수 있도록 하였으며 최종 결과치 분석에서도 시험데이터가 극치분포에 적합치 않은 경우에는 경고(warning)를 줄 수 있도록 하였다.

### IV. 결 론

프로그램 성능검토를 위해 표 2.1의 인위적 데이터[2]를 이용하여 수작업에 의한 그래픽방법과 컴퓨터 방법에 의한 모수추정 결과는 표 2.2와 같았다.

표 2.2의 수작업에 의한 그래픽방법의 결과치들은 실제 확률지를 사용하여 모수를 추정하도록 하여 얻은 추정치들의 평균데이터이다.

지금까지의 결과치 비교에서 나타난 공통적 사항은 역시 확률지 작도상의 오차가 가장 큰 원인으로 나타나 그래픽방법에 의한 모수추정의 경우 개인간에 많은 차이를 보였다. 그러나 컴퓨터에 의한 결과는 본 연구에서 의도했던 대로 정확하게 이론적인 수치를 얻을 수 있었다.

### V. 적용예

\*\*\*-\*\* \*-\*\*= L I M I T I N G P R O B L E M 1 \*\*\*-\*\* \*-\*\*=

INPUT DATA ;

5.00000	7.00000	8.50000	9.00000
10.00000	12.00000	13.00000	15.00000
16.00000	20.00000		

----- P R E L I M I N A R Y W O R K -----

N = 10 L = 10 INOF = 0  
 $F(i) = (i - .3) / (i + .4)$   $Y(i) = -\ln \ln 1/F(i)$

i	t	F(i)	Y(i)
1	5.00000	0.06731	-0.99269
2	7.00000	0.16346	-0.59398
3	8.50000	0.25962	-0.29903
4	9.00000	0.35577	-0.03292
5	10.00000	0.45192	0.23037
6	12.00000	0.54808	0.50860
7	13.00000	0.64423	0.82167
8	15.00000	0.74038	1.20202
9	16.00000	0.83654	1.72326
10	20.00000	0.93269	2.66385

\*\*\*\*\* R E S U L T \*\*\*\*\*

U = 9.402 ALPHA = 0.243  
 MU = 11.772 SIGMA^2 = 27.743

WARNING : THE FAILURE DATA MAY NOT SUIT LIMITING DISTRIBUTION  
 REGRESSION CURVE :  $Y = 0.002 X^2 + 0.186 X - 1.980$

RELIABILITY AT ORIGINAL DATA INPUT DATA	R(T)	RELIABILITY AT OTHER POINTS T I M E	R(T)
5.00000	0.93269	2.00000	0.99767
7.00000	0.83654	4.00000	0.97591
8.50000	0.74038	6.00000	0.89868
9.00000	0.64423	8.00000	0.75507
10.00000	0.54808	10.00000	0.57871
12.00000	0.45192	12.00000	0.41208
13.00000	0.35577	14.00000	0.27847
15.00000	0.25962	16.00000	0.18172
16.00000	0.16346	18.00000	0.11594
20.00000	0.06731	25.00000	0.02216

\*\*\*\*\*



==\*-\*= ==\*-\*= L I M I T I N G P R O B L E M 2 ==\*-\*= ==\*-\*=

INPUT DATA ;

1224.00000 1220.00000 1218.00000 1215.00000  
 1213.00000 1207.00000

----- P R E L I M I N A R Y W O R K -----

N = 6 L = 6 INOF = 0  
 $F(i) = (i - .3) / (i + .4)$   $Y(i) = -\ln \ln 1/F(i)$

i	t	F(i)	Y(i)
1	1224.00000	0.10938	-0.79434
2	1220.00000	0.26563	-0.28192
3	1218.00000	0.42188	0.14729
4	1215.00000	0.57813	0.60154
5	1213.00000	0.73437	1.17527
6	1207.00000	0.89062	2.15562

\*\*\*\*\* R E S U L T \*\*\*\*\*

U = 1218.970 ALPHA = 0.178  
 MU = 1215.734 SIGMA^2 = 51.696

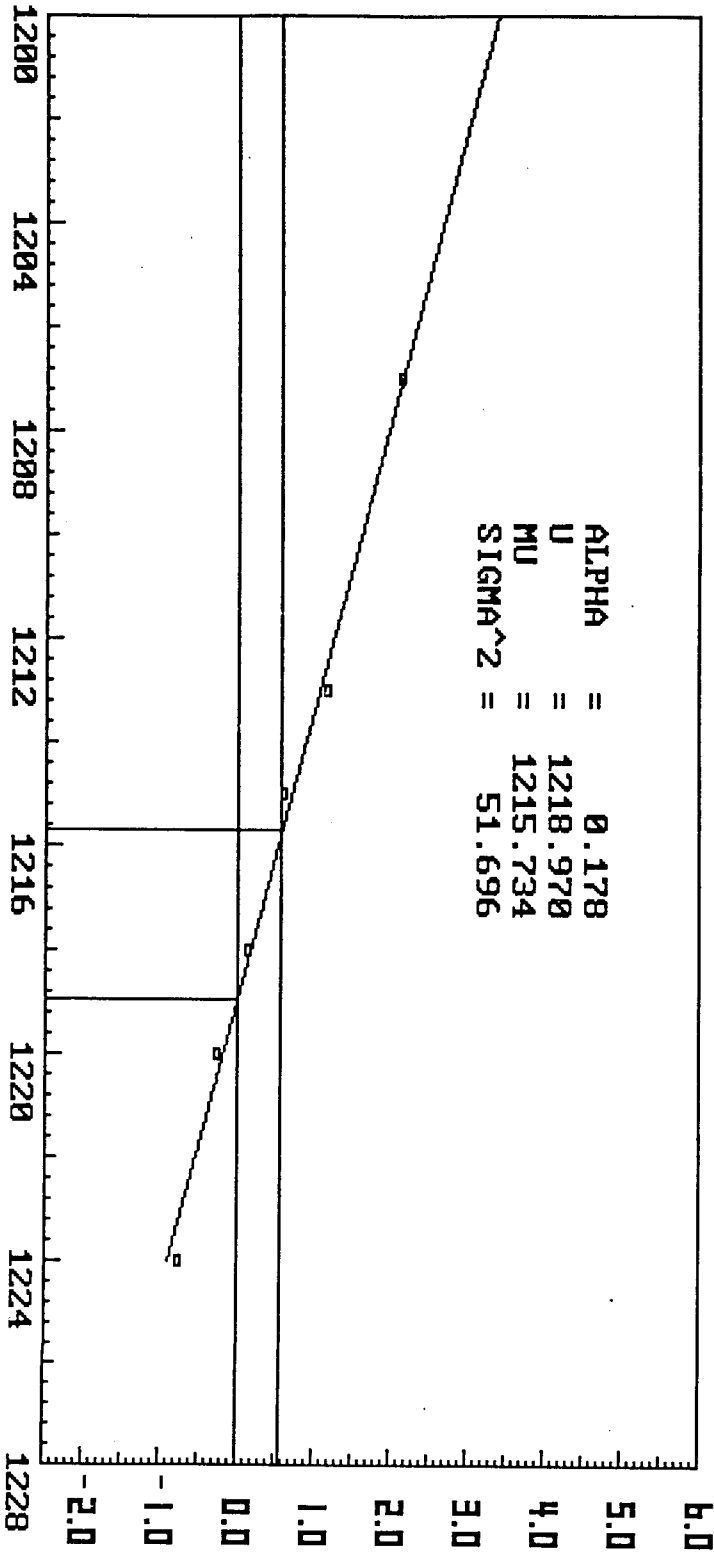
REGRESSION LINE : Y = 217.439000 - 0.178379 X

RELIABILITY AT ORIGINAL DATA		RELIABILITY AT OTHER POINTS	
INPUT DATA	R(T)	T I M E	R(T)
1224.00000	0.89063	1250.00000	1.00000
1220.00000	0.73438	1240.00000	1.00000
1218.00000	0.57813	1230.00000	0.99922
1215.00000	0.42188	1220.00000	0.69930
1213.00000	0.26563	1210.00000	0.18280
1207.00000	0.10938	1200.00000	0.03335

\*\*\*\*\*

>>>>>>>> L I M I T I N G P R O B A B I L I T Y P A P E R <<<<<<<<<  
REGRESSION LINE :  $y = 217.439000 - 0.178379 x$  -lnln 1/F(t)

ALPHA = 0.178  
U = 1218.970  
MU = 1215.734  
SIGMA^2 = 51.696



==\*-== ==\*-== L I M I T I N G P R O B L E M 3 ==\*-== ==\*-==

INPUT DATA ;

116.00000 73.00000 46.00000 30.00000  
 19.00000 7.00000

----- P R E L I M I N A R Y W O R K -----

N = 6 L = 6 INOF = 0  
 F(i)=i/(N+1) Y(i)=-lnln 1/F(i)

i	t	F(i)	Y(i)
1	116.00000	0.14286	-0.66573
2	73.00000	0.28571	-0.22535
3	46.00000	0.42857	0.16570
4	30.00000	0.57143	0.58051
5	19.00000	0.71429	1.08924
6	7.00000	0.85714	1.86983

\*\*\*\*\* R E S U L T \*\*\*\*\*

U = 70.666 ALPHA = 0.021  
 MU = 43.387 SIGMA^2 = 3673.868

WARNING : THE FAILURE DATA MAY NOT SUIT LIMITING DISTRIBUTION  
 REGRESSION CURVE : Y = 0.000 X^2 - 0.052 X - 2.073

RELIABILITY AT ORIGINAL DATA		RELIABILITY AT OTHER POINTS	
INPUT DATA	R(T)	T I M E	R(T)
116.00000	0.85714	10.00000	0.24196
73.00000	0.71429	20.00000	0.28986
46.00000	0.57143	40.00000	0.40704
30.00000	0.42857	80.00000	0.70428
19.00000	0.28571	90.00000	0.77809
7.00000	0.14286	120.00000	0.94159

\*\*\*\*\*





## 參 考 文 獻

- [1] 市田嵩, 鈴木和幸, 신뢰성의 분포와 통계, 한국공업표준협회 역, 김 재주 감수, 한국공업표준협회, 1992
- [2] 鹽見弘 외 3인, 신뢰성에서 확률지의 사용, 한국공업표준협회 역, 김 재주 감수, 한국공업표준협회, 1992
- [3] Fisher, R. A. and L. H. C. Tippett, *Limiting Forms of the Frequency Distribution of Large or Smallest Number of a Sample*, Proc. Cambridge Philos. Soc., 24, Pt. 2, 1928
- [4] Gumbel, E. J., 극치통계학, 河田龍夫 외 역, 생산기술센터, 1978
- [5] Ireson, W. G., *Reliability Handbook*, McGrawhill Book Company, 1966
- [6] Kimball, B. J., On the Choice of Plotting Positions on Probability Paper, *Journal of the American Statistical Association*, Vol 55, 1960
- [7] Lawless, J. F. and Mann, N. R., Test for Homogeneity for Extreme Value Scale Parameter, *Commun. Stat.*, Vol. A5, 1976
- [8] Maurice G. Kendall, Alan Stuart, *The Advanced Theory of Statistics : Distribution Theory*, 3rd ed., Vol. 1, Hafner Publishing Company, New York, 1969
- [9] \_\_\_\_\_, *The Advanced Theory of Statistics : Inference and Relationship*, 2nd ed., Vol. 2, Hafner Publishing Company, New York, 1967