고장데이타 수집 및 신뢰성 해석

張 重 淳
亞洲大學校 産業工學科

Ⅰ. 서 언

최근들어 국내의 많은 기업은 제품의 국제경쟁력을 확보하기 위하여 그 제품의 신뢰성을 적극적으로 높여야 한다는 인식을 하게 되어 신뢰성 시험설비를 확충하는 등 이 분야의 투자를 늘려나가려 하고 있다. 이는 그 동안은 외국과의 합작 등으로 인하여 기업의 관심은 좋은 제품의 설계보다는 어떻게 하면 모든 품질규격을 만족하는 양질의 제품을 제조할 것인가에 있었던 반면, 요즈음은 여러 여건의 변화로 인하여 국내 각 기업이 제품설계와 제조를 자체적으로 해결하여야 하는 경우가 많아지게 되었는 바, 적극적인 품질보증을 위하여는 고신뢰성 제품의 생산이 절대적으로 필요하다는 것을 인식하였기 때문이라고 볼 수 있다.

고신뢰성 제품을 생산하기 위하여는 제품 설계시 그 설계 대안에 의해 제조된 제품에서 발생할 수 있는 고장들을 미리 예측하여 제거하고 그 결과를 각종의 신뢰성 시험을 통하여 확인하여야 하는데, 이와 같은 과정에는 필연적으로 많은 양의 정보가 필요하고 또 많은 양의 새로운 정보가 만들어 지게 된다. 예를들어 제품 설계시에나 공정의 설계시에는 시장정보, 환경정보, 경쟁품의 기능·성능에 관한 정보, 기존 모델의 고장에 관한 정보, A/S 정보 등등의 정보가 필요할 것이다. 이와 같은 정보들은 계량화할 수 있는 것도 있고 그렇지 못한 것도 있을 수 있다. 그러나 어느 경우라도 이들을 효율적으로 관리할 수 있는 데이타베이스를 갖추고 있어야, 해당 제품의 설계·생산시에는 물론 유사제품의 설계시에도 활용할 수 있을 것이다.

이 글에서는 이들 정보중에서 제품의 신뢰성 데이타를 분석하는 방법을 다루고자 한다. 신뢰성 데이타란 시스템이나 제품, 또는 부품 (이하 아이템 이라고 한다.)의 고장의 형태와 고장이 발생하였을 때까지의 작동시간 즉 수명에 관한 정보를 말하는 것이다. 신뢰성 데이타는 해당 아이템이 신뢰도 목표치를 만족하고 있는가를 판단하거나, 아이템의 개선을 위하여 결함을 찾아내고, 또 아이템들의 비교평가 또는 신뢰도 예측을 위한 고장이력을 마련하기 위하여 사용되어 진다. 신뢰성 데이타는 또한 시스템의 정비, 로지스틱스, 정상운전 등에 필요한 정보도 제공한다. 신뢰성 데이타는 신뢰도 시험데이타, 고장해석 및 분석데이타, A/S 데이타, 필드모니터(field monitor) 데이타, 협력업체 데이타, 신뢰성 데이타뱅크 등에서 나올 수 있다. 그런데 이와 같은 신뢰성 데이타는 반드시 고장현상을 사실적으로 또 정확하게 표현한 것이어야 한다. 왜냐하면 이 데이타들이 완벽하지 못하면 그들로 부터 얻어진 결론은 당연히 부정확할 수 밖에 없기 때문이다.

이 글에서는 신뢰성 데이타중 고장시간 데이타를 분석하는 방법만을 간략히 소개하고자 한다. 고장시간데이타 분석방법은 그동안 많은 통계학자들이 연구를 하여 여러 권의 책도 발간되어 있는데, 여기에서는 이들 중 가장 손쉽게 활용되어질 수 있는 것들만을 소개하고자 한다.

Ⅱ. 고장분포함수

고장시간데이타를 분석하는 방법에는 크게 그래프에 의한 방법과 통계적인 방법이 있는데, 어느 경우

나 첫째, 아이템의 고장시간의 확률분포함수 (probability distribution function)와 그 모수 (parameter)를 추정하고, 둘째, 평균수명, 고장률 등과 신뢰도 지표를 추정하며, 셋째, 모수나 신뢰도 지표등에 대한 신뢰구간을 구하기 위하여 사용된다. 이 절에서는 먼저 신뢰성분석의 기초가 되는 개념과 정의에 대하여 간단히 알아보고자 한다.

T를 어떤 한 아이템의 수명을 나타내는 확률변수라고 하자. T의 누적 분포함수

$$F(t) = \Pr\{T \leq t\}, \quad t > 0$$

는 아이템이 t시간 이전에 고장날 확률을 나타내고 있는데, 이 F(t)를 고장분포함수라고 부른다. 아이템이 t시간 이상을 고장없이 작동할 확률을 시간 t의 함수로 나타낸 것을 신뢰도함수(reliability function)라 한다. 즉 신뢰도 함수는

$$R(t) = \Pr\{T > t\} = 1 - F(t)$$

이 된다.

고장시간 T는 경우에 따라서는 정수값을 취하는 경우도 있으나, 대부분은 실수값을 갖는다고 볼 수 있으므로, 이 글에서는 후자의 경우만을 다루고자 한다. 이 경우 F(t)의 도함수

$$f(t) = dF(t)/dt$$

는 고장확률 밀도함수라고 한다. 이는 고장시간 데이타의 히스토그램에 해당하는 함수이다.

고장분포를 표현하는 또 다른 방법은 그 분포의 고장률함수(hazard rate function 또는 failure rate function)를 고려하는 것이다. 고장률함수 h(t)는 t 시간까지 고장나지 않은 아이템이 t시각 직후 고장날 조건부확률의 변화율을 말하는 것인데,

$$\begin{aligned} h(t) &= \lim_{\Delta t \downarrow 0} \Pr\{T \leq t + \Delta t \mid T > t\}/\Delta t \\ &= f(t)/(1 - F(t)) \qquad (1) \\ &= f(t)/R(t) \end{aligned}$$

과 같이 구해진다. 이 식으로부터

$$\begin{aligned} &R(t) = \exp(-H(t)), \\ &f(t) = h(t)\ \exp(-H(t)), \qquad (2) \\ &\text{단, } H(t) = \int h(u)\,du \end{aligned}$$

등의 관계를 얻을 수 있다. 여기서 H(t)는 누적고장률함수라고 한다.

고장률함수 h(t)는 고장시간데이타가 어떤 분포를 따르는가에 따라서 여러가지 형태가 있을 수 있으며, 전형적인 몇가지 유형을 보면 그림1과 같이

(a) 증가형 고장률함수(increasing failure rate function: IFR)
(b) 감소형 고장률함수(decreasing failure rate function: DFR)
(c) 욕조형 고장률함수(bath-tub failure function)
(d) 상수형 고장률함수(constant failure rate function: CFR)

등으로 나눌 수 있다. 그림1에서 (a)는 IFR을 나타내는데, 이것은 사용기간에 비례하여 고장이 많이 발생하는 아이템의 고장률에 해당하며, (b)는 DFR의 경우로 오래 사용한 것일수록 고장날 가능성이 줄어드는 아이템에 해당된다. (c)는 욕조형 고장률함수로서 초기에는 제조미스 또는 오용등의 원인으로 인하여 고장이 많이 발생하다가 시간이 지남에 따라 고장률이 감소하여 일정한 고장률을 가지며, 제품을 오래 사용하면 마모 등의 원인으로 인하여 다시 고장이 많이 발생하는 아이템에 해당된다. (d)는 사용기간에 관계없이 고장날 가능성이 일정한 경우이다.

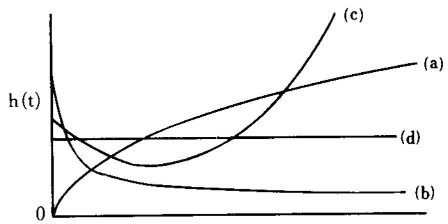


그림 1. 고장률함수의 유형

아이템의 신뢰도지표로서 많이 사용되는 평균수명은 그 아이템이 고장났을 때 수리가 가능한가의 여부에 따라서 가능하면 MTBF(mean time between filures)라고 하고, 불가능하면 MTTF(mean time to failure)라고 한다. MTBF와 MTTF는

$$\text{MTBF(MTTF)} = \int t f(t)\,dt$$

로 정의된다.

신뢰성분석에 많이 활용되는 분포로는 지수분포(exponential distribution), 와이블분포(Weibull distribution), 정규분포(normal distribution), 대수정규분포

(lognormal distribution), 감마분포(gamma distribution), 극치분포(extreme value distribution)등이 있다.
지수분포는 고장확률밀도함수가

$$f(t)=\lambda \cdot \exp(-\lambda t), \quad t>0 \tag{3}$$

과 같이 표현되어지는 경우를 말한다. 지수 분포의 경우에는

$$F(t)=1-\exp(-\lambda t) \tag{4}$$

이므로,

$$h(t)=\lambda$$

로서 시간에 따라 변화하지 않는 상수형 고장률함수가 된다. 고장률함수가 상수라는 것은 과거 이력에 관계없이 작동중 어느 순간에서나 고장날 가능성이 동일하다는 것을 의미하는 것으로, 이와 같은 형태의 고장은 랜덤하게 발생하는 쇼크(shock)를 이기지 못해 나가는 경우에 해당된다. 대부분의 전자부품의 고장시간은 이 지수분포로써 잘 표현될 수 있는 것으로 알려져 있다.

와이블분포는

$$f(t)=\alpha\beta \ t^{\beta-1} \exp(-\alpha t^{\beta}), \quad t>0$$
$$F(t)=1-\exp(-\alpha t^{\beta}) \tag{5}$$
$$h(t)=\alpha\beta \ t^{\beta-1}$$

과 같은 확률분포를 갖는 것이다. 이 분포는 특히 기계부품의 고장시간을 표현하는데 적합한 것으로 알려져 있다. 특히 $\beta=1$인 경우에는 지수분포와 같은 형태가 되며, $\beta>1$이면 고장률함수 $h(t)$는 증가형 고장률함수가 되며, $\beta<1$인 경우에는 감소형 고장률함수가 된다.

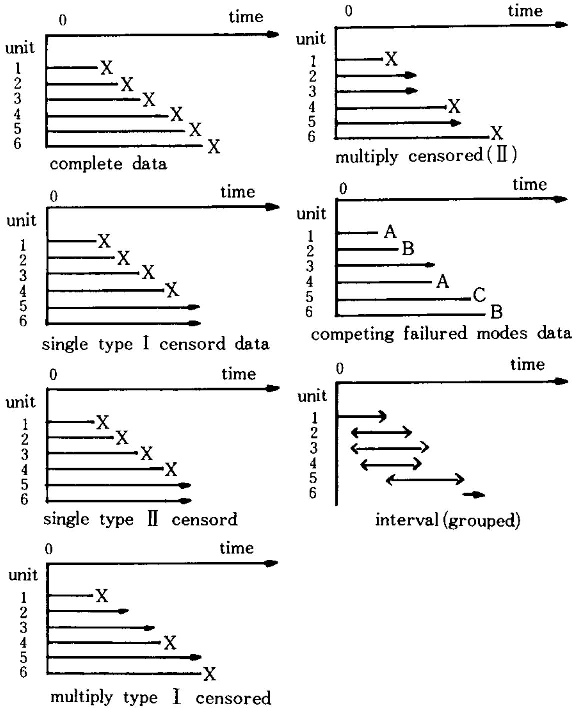
한편 정규분포, 대수정규분포, 감마분포, 극치분포등과 이들의 혼합형에 대하여는 참고문헌 [2]에 잘 설명되어 있다.

Ⅲ. 고장시간 데이타의 유형

고장시간데이타는 앞에서 살펴본 바와 같이 신뢰성시험이나 A/S, 또는 모니터 등등을 통하여 얻어질 수 있다. 그런데 이러한 데이타들은 모든 아이템의 고장시간을 전부 파악하였는가의 여부에 따라 완전데이타(complete data)와 관측중단데이타 (censored data)로 구분되어 진다.

관측중단데이타란 모든 아이템의 고장시간을 일부만을 관측한 경우의 데이타이다. 관측중단데이타는 관측중단방법에 따라 정시종결 (type Ⅰ censored) 데이타와 정수종결 (type Ⅱ censored) 데이타등으로 나누어진다. 정시종결 데이타는 예를 들어 신뢰성시험시 일정시간이 되면 아이템이 모두 고장나지 않더라도 시험을 중단하는 경우를 말하고, 정수종결 데이타는 일정갯수의 아이템이 고장나면 관측을 중단하는 경우를 말한다. 따라서 만일 신뢰성시험시 500시간이 되면 관측중단을 할 경우에는 정시종결 데이타이고, 20개의 아이템을 시험하여 그 중 10개가 고장나 관측을 중단하면 정수종결데이타가 된다.

또 이와 같은 관측중단 시점이 하나만 있는 경우에는 단순종결데이타(singly censored data)라고 하고, 여러개 있는 경우에는 다종종결데이타(multiply censored data)라고 한다. 다종종결데이타는 대부분의 A/S 데이타와 같은 필드데이타에서 발견될 수 있다. 왜냐하면 필드데이타의 경우에는 데이타분



※ Failure time x, running time→, Failure occurred earlier←

그림 2. 데이타의 유형별 도시

석시점에서 볼 때는 제품들이 사용되기 시작한 시간이 모두 다르기 때문이다.

어떤 경우에는 고장발생시의 고장모드가 여러개 있을 수도 있다. 이와 같은 경우를 다종고장모드 데이타 (competing failure mode data)라고 부른다. 예를 들어 모타의 고장에는 소손도 있을 수 있고 단선도 있을 수 있는 바, 이들 모두 모터를 정지시키는 고장모드들이다. 그런데 이와 같은 고장 모드들을 분리하지 않고 모두 합하여 고장시간을 기록하였다면 이를 다종고장모드데이타가 되는 것이다.

어떤 경우에는 또한 아이템을 점검 (inspection) 하여 보아야만 고장 발생여부를 확인할 수 있는데, 이러한 경우에는 아이템의 고장시간을 정확하게 파악할 수 없고, 다만 고장이 언제부터 언제 사이에 발생하였을 것이라는 정보만을 얻을 수 있다. 이러한 데이타를 구간 또는 묶음데이타(interval or grouped data)라고 한다.

고장시간데이타는 이와 같은 여러가지 형태로 구분되어질 수 있는데, 반드시 각 경우에 알맞는 정확한 분석방법에 따라 처리되어야 한다. 그림2는 고장데이타를 그림으로 표시한 것이다.

Ⅳ. 그래프에 의한 분석법

고장시간을 분석하는 방법은 먼저 그 데이타의 확률분포를 알아내는 것부터 시작되어야 한다. 그런데 고장시간데이타의 수가 많으면 히스토그램을 작성하여 확률분포를 찾는 것이 가장 용이한 방법이다. 그러나 고장데이타는 그 수도 적은 것이 보통이고, 또 관측중단된 데이타도 혼입되어 있으므로 확률용지(probability paper)나 고장률용지(hazard rate paper) 또는 총 시험시간(total time on test:TTT) 타점법 등을 이용하여 분포를 알아내는 것이 일반적인 방법이다.

1. 확률용지분석법

확률용지는 분석하고자 하는 아이템의 분포함수가 가정한 분포를 따른다면, 얻어진 데이타를 확률용지상에 타점하였을 때 그들이 거의 직선에가깝도록 가로축과 세로축의 스케일(scale)을 조정하여 놓은 용지로써, 타점된 데이타들을 연결하였을 때 그들이 직선에서 크게 벗어나지 않는다면 가정한 분포를 따른다고 결론을 내릴 수 있다.

확률용지는 각 확률분포에 따라 다르게 만들어 진다. 예를 들어 지수분포에서는 식(4)로부터

$$t = -\ln(1-F(t))/\lambda \qquad (6)$$

의 관계를 얻을 수 있는데, 따라서 t와 $\ln(1-F(t))$는 선형관계에 있음을 알 수 있다. 이를 이용하여 지수확률용지는 가로축에 고장시간 t를, 세로축에는 $F(t)$를 타점하여 $(t, \ln(1-F(t)))$가 직선을 이룰 수 있도록 스케일이 조정되어 있다.

한편 와이블분포에서는 식(5)로부터

$$\ln\ln\left(\frac{1}{1-F(t)}\right) = \ln\alpha + \beta\cdot\ln t \qquad (7)$$

의 관계를 얻을 수 있다. 이를 이용하여 만들어진 와이블확률지도 지수확률 용지와 마찬가지로 가로축에 고장시간 t를, 세로축에는 $F(t)$를 타점하되, 다만 $(\ln t, \ln\ln(1/(1-F(t))))$가 직선을 이룰 수 있도록 스케일이 조정되어 있다는 점이 다를 뿐이다.

다른 분포들도 이와 비슷한 방법으로 확률용지를 설계할 수 있다. (참고문헌[2]참조) 그러면 지금부터는 실제로 고장시간 데이타가 주어졌을 때, 확률용지에 타점하는 방법에 대하여 알아보기로 한다. 지금 n개의 아이템을 시험하여 고장시간이 $t_1, \cdots, t_n$으로 얻어졌다고 가정하자. 확률용지에 이 데이타들을 타점하기 위하여는 먼저 누적분포함수 $F(t)$에 대한 추정치를 구하여야 하는데 대개 다음과 같은 3가지 방법이 많이 쓰인다.

$$\text{(a) 평균 랭크법}: \hat{F}(t) = \frac{r}{n+1}$$

$$\text{(b) 메디안 랭크법}: \hat{F}(t) = \frac{(r-0.3)}{(n+0.4)} \qquad (8)$$

$$\text{(c) 경험적 방법}: \hat{F}(t) = \frac{r}{n}$$

단 여기서 $\hat{F}(t)$는 $F(t)$의 추정치이며, r은 t시간 이전에 고장난 아이템의 수이다. r의 값을 구하기 위하여는 $t_1, t_2, \cdots t_n$은 크기 순서대로 정리하여 순위(rank)를 매기는 것이 편리하며, 타점은 각 고장시간 $t_1, t_2, \cdots t_n$의 위치에서만 하면 된다.

예 1. 어떤 전자부품 20개를 수명시험하여 다음과 같은 데이타를 얻었다.

4, 19, 59, 126, 42, 22, 49, 106, 25, 63
80, 35, 73, 160, 39, 11, 7, 16, 94, 32

이 데이타를 평균랭크법을 이용하여 와이블 확률지에 타점하여 보면 다음과 같다.

순 위	1	2	3	4	5	6	7	8	9	10
고장시간	4	7	11	16	19	22	25	32	35	39
$\hat{F}(t)$ (%)	4.8	9.5	14.3	19.0	23.8	28.6	33.3	38.1	42.9	47.6

순 위	11	12	13	14	15	16	17	18	19	20
고장시간	42	49	59	63	73	80	94	106	126	160
$\hat{F}(t)$ (%)	52.4	57.1	61.9	66.7	71.4	76.2	81.0	85.7	90.5	95.2

이 데이타를 와이블 확률지에 타점하기 위하여는 고장시간의 스케일을 적절히 조정하여야 한다. 이와같이 하여 타점된 점들을 보면 대개 직선관계를 갖고 있는 것으로 판단되며, 따라서 이 고장시간의 분포는 와이블 분포라고 간주할 수 있다. 한편 와이블 분포의 모수인 β의 값을 구해보면 (5절 참조), $\beta \fallingdotseq 1.0$인 것을 알 수 있다. $\beta \fallingdotseq 1.0$이라는 것은 이 고장데이타의 분포는 지수분포라고 간주하여도 좋다는 것을 의미한다.

고장시간 데이타에 관측중단된 데이타가 혼입되어 있을 때에는 관측 중단된 데이타도 고장시간과 같이 간주하여 순위와 F(t)의 추정치를 구하지만, 타점은 고장난 시간이 관측될 경우에만 행하고 관측중단된 데이타는 타점하지 않는다는 점만 다르다.

2. 고장률 용지 분석법

고장률용지도 확률용지와 마찬가지로 고장시간 분포를 알아내는데 쉽게 사용할 수 있는 그래프이다. 고장률용지의 설계이론을 간단히 살펴보자. 먼저 고장시간 분포가 지수분포인 경우에는 고장률함수가 $h(t)=\lambda$이므로, 누적 고장률 함수는

$$H(t)=\lambda t \qquad (9)$$

이 된다. 이 식으로 부터 지수분포의 경우에는 t와 H(t)는 선형관계에 있음을 알 수 있다. 한편 와이블 분포에서는

$$H(t)=\alpha t^{\beta}$$

이므로

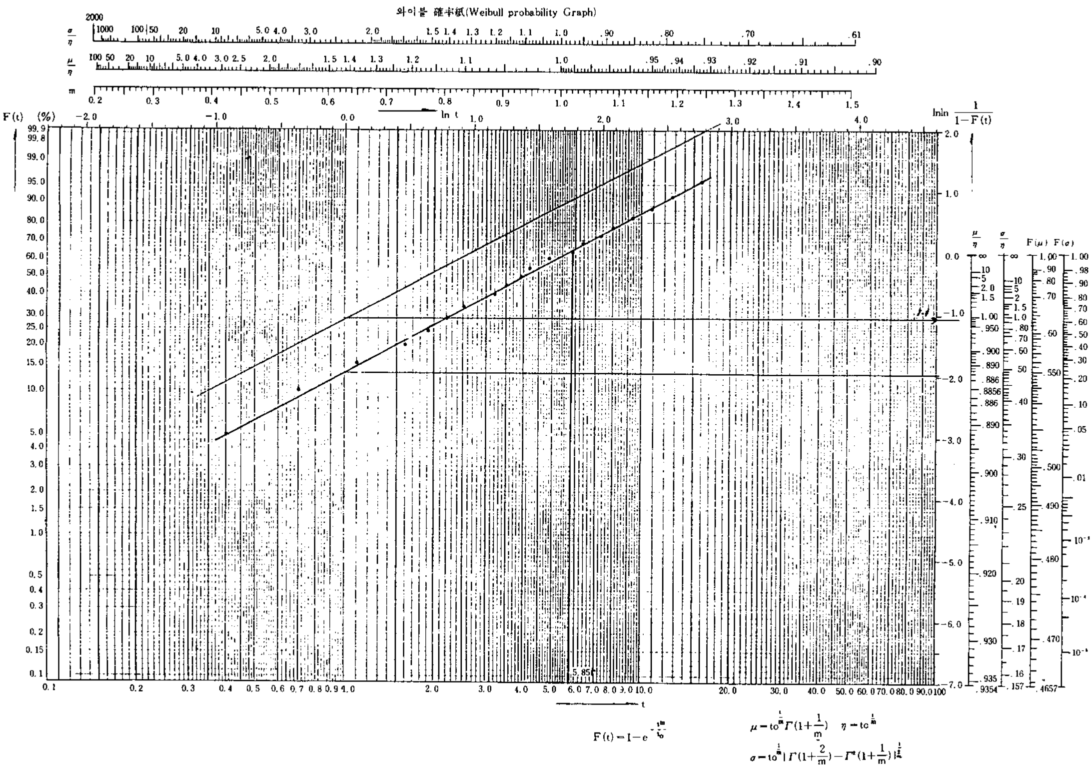


그림 3. 와이블 확률지 타점

$$\ln H(t) = \ln \alpha + \beta \ln t \qquad (10)$$

이다. 따라서 와이블 분포에서는 $\ln t$와 $\ln H(t)$가 선형관계에 있음을 알 수 있다. 고장률용지는 이와 같은 누적고장률 함수의 성질을 이용하여 고장시간 분포를 알아내자는 것이다.

고장률용지를 사용하기 위하여는 누적고장률 함수의 측정치를 구하여야 하는데, 일반적으로 다음과 같은 방법이 사용된다. 고장률 함수 $h(t)$의 정의에 의하여 $h(t)$의 추정은 t시간까지 고장나지 않았던 아이템 중에서 t시각 직후 몇개나 고장 났는가 하는 비율로 추정되며, $H(t)$는 t이전의 고장률 함수의 추정치를 모두 합하여 구한다.

예 2. 예1의 데이타를 지수분포의 고장률용지에 타점하여 보면 다음과 같다.

고장시간	4	7	11	16	19	22	25	32	35	39
$\hat{h}(t)$	1/20	1/19	1/18	1/17	1/16	1/15	1/14	1/13	1/12	1/11
$\hat{H}(t)$ (%)	3	10.3	15.8	21.7	28.0	34.6	41.8	49.5	57.8	66.9

고장시간	42	49	59	63	73	80	94	106	126	160
$h(t)$	1/10.	1/9	1/8	1/7	1/6	1/5	1/4	1/3	1/2	1/1
$\hat{H}(t)$ (%)	76.9	88	100.5	114.8	131.5	151.5	176.5	209.8	259.8	357.8

여기서 $h(t)$와 $\hat{H}(t)$를 구하는 방법을 살펴보면, 예를 들어 두번째 고장시간인 7시간에서는 작동하고 있던 아이템이 19개 이었는데 그 중 한개가 고장났으므로 $\hat{h}(t)$은 1/19이고, 따라서 누적고장률 $\hat{H}(t) = 1/20 + 1/19 \fallingdotseq 0.103$임을 알 수 있다.

다음 그림4는 이 고장률 용지를 컴퓨터로 작성한 것이다. 이 고장률지를 보면 타점된 점들이 거의직선관계에 있어, 고장분포함수는 지수분포로 가정하여도 무방함을 알 수 있다. 관측중단데이타가 혼입되어 있는 경우의 고장률 용지의 작성은 확률용지의 경우와 같은 방법으로 처리한다.

3. 총시험시간 타점법

총 시험시간 타점법은 고장률함수가 IFR인가 DFR인가 CFR인가를 알아볼 수 있는 방법이다. 만일 n개의 아이템을 수명시험하여 제일 먼저 고장난 시간 부터 차례로 $t_1, t_2, \cdots, t_n$이라고 한다면 i번째 고장시간 t_i까지 관찰된 총 시험시간은

$$T(t_i) = nt_1 + (n-1)(t_2 - t_1) + \cdots + (n-i+1)(t_i - t_{i-1}) \qquad (12)$$

이 된다. 즉 이는 t_1까지는 n개의 아이템이 작동하고 있었고, t_1부터 t_2까지는 $(n-1)$개의 아이템이 작동하고 있었으며, t_{i-1}에서 t_i까지는 $(n-i+1)$개의 아이템이 작동하였으므로, 그 시간을 모두 합한 것이다.

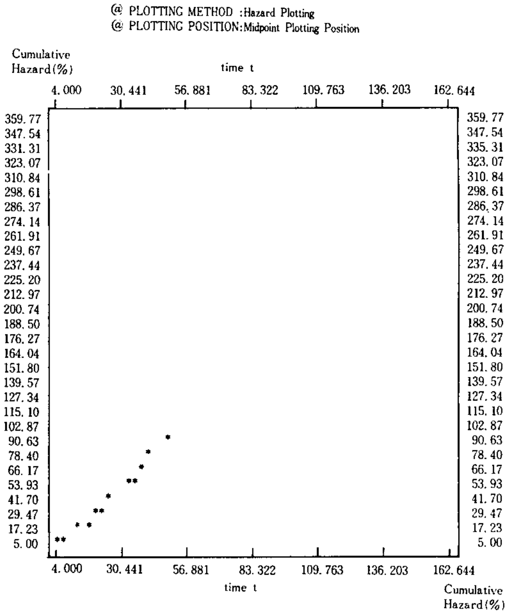


그림 4. 지수고장률 용지 타점

〈T. T. T. Plot〉

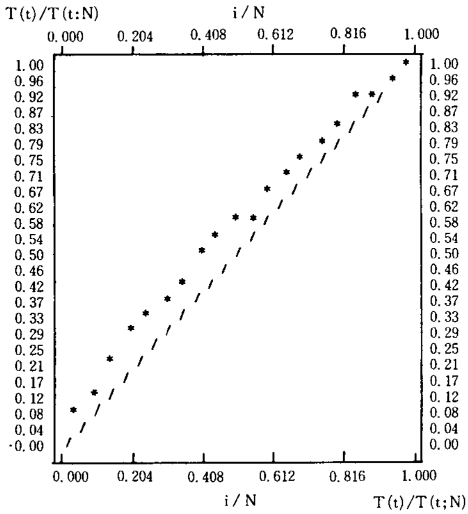


그림 5. 총 시간타점

총 시간 타점법은 $T(t_i)/T(t_n)$의 값을 i/n에 대하여 나타낸 것이다. 모든 관찰된 데이타를 이 방법에

의해 타점해 이 점들이 직선을 이루면 고장률이 CFR이고, 볼록한 모양을 이루면 IFR이며, 오목한 모양을 이루면 DFR임을 알 수 있다.

예 3. 그림5는 예1의 데이타에 대한 총 시간 타점 그래프이다. 타점된 모양을 보면 점들이 대각선 쪽으로 거의 직선이라 볼 수 있어 고장률 함수는 CFR이라 볼 수 있다.

V. 통계적 방법

앞절에서는 고장시간 데이타를 그래프를 이용하여 분석하는 방법에 대하여 알아보았다. 그래프에 의한 방법은 고장분포 함수를 찾아내는데 매우 용이하게 활용될 수 있기는 하지만, 많은 경우 판정기준이 직관에 의존하고 있어서, 타점이 이루어지고 나면 분포에 대한 적합성 검정(goodness of fit test)등의 통계적 절차를 거치는 것이 바람직하다.

이 절에서는 고장시간 데이타들을 통계적으로 처리하여 분포의 모수나 신뢰도지표등에 대한 점추정 또는 구간 추정을 행하는 방법에 대하여 알아보고자 한다. 이와 같은 추정을 위하여는 고장 분포함수에 대한 적절한 형태를 가정하고 분석할 수도 있고, 그러한 가정없이 행할 수도 있다. 전자와 같은 통계분석을 모수적방법 (parametric method)라고 하고, 후자와 같은 방법을 비모수적방법(nonparametric method)라고 한다. 이 절에서는 먼저 모수적방법중에서 지수분포를 따르는 데이타의 분석법 부터 알아본다.

1. 지수분포 데이타의 분석

지수분포의 고장분포를 갖는 n개의 아이템을 수명시험하여 $t_1, t_2, \cdots, t_n$과 같은 데이타를 얻었다고 하자, 이 경우 고장률 λ를 추정하는 방법은 흔히 최우추정법(maximum lihelihood estimation)을 많이 이용한다. 최우 추정법은 표본 즉 데이타의 결합 확률밀도함수(joint p. d. f)를 최대로 하는 모수의 값을 추정치로 하자는 것이다. 따라서 λ의 추정치는

$$L(\lambda) = \prod_{i=1}^{n} \lambda \ \exp(-\lambda t_i)$$
$$= \lambda^n \exp(-\lambda \sum_{i=1}^{n} t_i)$$

를 최대로 하는 값인데, 미분하여 풀어보면

$$\hat{\lambda} = \frac{n}{\sum t_i} \tag{13}$$

임을 알 수 있다. 여기서 $\theta \equiv$MTBF라고 하면 θ의 추정치는

$$\hat{\theta} = 1/\hat{\lambda} \tag{14}$$

과 같이 됨을 알 수 있다.

다음으로 관측 중단된 데이타의 경우를 알아보자. 지금 위의 수명시험을 t_c시간에서 관측중단하는 정시종결방식을 채택하였다고 가정하자. 만일 t_c 시간까지 고장난 아이템의 수가 r개이라면, (n−r)개의 아이템은 t_c시간까지 고장이 발생하지 않았으므로, 표본의 결합 확률밀도 함수는

$$L(\lambda) = \prod_{i=1}^{n} \lambda \exp(-\lambda t_i) \prod_{i=r+1}^{n} \exp(-\lambda t_c)$$
$$= \lambda^r \exp\{-\lambda (\sum_{i=1}^{r} t_i + (n-r) t_c)\}$$

임을 알 수 있다. 따라서 고장률 λ의 최우 추정치는

$$\hat{\lambda} = \frac{r}{\sum_{1}^{r} t_i + (n-r) t_c} \tag{15}$$

가 된다.

또 만일 수명시험을 r개의 아이템이 고장날때까지만 관측하는 정수종결 방식을 택하였다면

$$L(\lambda) \lambda^r \exp\{-\lambda (\sum_{i=1}^{r} t_i + (n-r) t_r)\}$$

이 됨을 알 수 있는 바, 따라서

$$\hat{\lambda} = \frac{r}{\sum_{1}^{r} t_i + (n-r) t_r} \tag{16}$$

이 된다. 그런데 식(13), (15), (16)의 분모는 모두 총시험시간을 나타내고 있고, 분자는 고장난 아이템의 수를 나타내고 있음을 알 수 있다. 따라서 이들을 종합하면 어느 경우에서나

$$\hat{\lambda} = \frac{\text{총 고장수}}{\text{총 시험시간}} \tag{17}$$

가 되고

$$\hat{\theta} = \frac{\text{총 시험시간}}{\text{총 고장수}} \tag{18}$$

와 같이 구할 수 있음을 알 수 있다.

또 아이템이 t^*시간의 고장없이 작동될 신뢰도는

$$\hat{R}(t^*) = \exp(-\hat{\lambda} t^*)$$

와 같이 구할 수 있다.

예 4. 지수분포의 수명분포를 갖는다고 생각되는 20개의 아이템을 수명시험하여 100시간에서 관측중단 하였다. 관측된 고장시간은 10, 16, 17, 25, 31, 46, 65시간 이었다. 고장률과 MTBF에 대한 추정치는?

(해) 이 경우 고장 갯수는 r=7이고, 총 시험시간은

$$\sum_{1}^{7} t_i + (20-7) * 100 = 1510$$

이므로, 고장률에 대한 추정치는

$$\hat{\lambda} = 7/1510 = 4.6 * 10^{-3} \text{(회/시간)}$$

이고,

$$\hat{\theta} = 215.7 \text{ (시간)}$$

임을 알 수 있다. 이 아이템이 100시간 이상을 고장없이 작동할 신뢰도는

$$\hat{R}(100) = \exp(-0.46) = 0.63128$$

과 같이 추정할 수 있다.

지수분포의 경우 고장률 또는 MTBF에 대한 구간추정은 관측중단 방법에 따라 달라진다. 먼저 정수 종결방식의 경우 MTBF에 대한 $(1-\alpha) * 100\%$의 신뢰구간은

$$\left[\frac{2T}{\chi^2_{2r}(1-\alpha/2)} \frac{2T}{\chi^2_{2r}(\alpha/2)}\right] \tag{19}$$

이다. 단 여기서 T는 총 시험시간이며, $\chi^2(p)$는 자유도 ϕ를 갖는 x^2분포에서 이 값보다 클 확률이 p인 백분위수이다. 고장률 λ의 신뢰 구간은 식(19)의 역수를 취하면 된다.

예 5. 지수분포의 고장분포를 갖는 아이템 12개를 수명시험하여 8개가 고장났을 때 관측 중단하였다. 관측된 고장시간은 다음과 같다.

31, 58, 157, 185, 300, 470, 497, 673

여기서 부터 MTBF에 대한 95% 신뢰구간을 구하면 다음과 같다. 먼저 총 시험시간 T=5063이다. 또 r=8이므로 χ^2 분포표에서

$$\chi^2_{16}(0.025) = 28.8 \qquad \chi^2_{16}(0.975) \fallingdotseq 6.91$$

을 얻을 수 있다. 따라서 MTBF의 95% 신뢰구간은

$$\left[\frac{2*5063}{28.8}, \frac{2*5063}{6.91}\right] = [351.6, 1469.4]$$

이다. 따라서 고장률의 95% 신뢰구간은

$$[1/1469.4,\ 1/351.6] = [0.00068,\ 0.00284]$$

이 된다.

완전데이타는 정수 종결 방식에서 r=n인 경우로 생각할 수 있으므로 MTBF에 대한 구간추정은 식(19)와 같이 하면 된다. 한편 정시 종결데이타의 경우에는 MTBF에 대한 구간추정은 근사적으로 다음과 같이 구한다.

$$\left[\frac{2T}{\chi^2_{2(r+1)}(\alpha/2)} \frac{2T}{\chi^2_{2r}(1-\alpha/2)}\right] \tag{20}$$

단, 여기서 T는 총시험시간이고, r은 고장난 아이템의 수이다. 만일 r=0이면 위 식은 사용할 수 없다.

2. 와이블 분포의 고장데이타 분석

와이블 분포를 따르는 경우의 고장시간 데이타의 분석은 지수분포의 경우보다 복잡하게 된다. 먼저 완전데이타의 경우를 살펴보면 와이블 분포의 모수 α와 β에 대한 최우추정치들은

$$L(\alpha,\beta) = \prod_{i=1}^{n} (\alpha\beta t_i^{\beta-1} \exp(-\alpha t_i^{\beta}))$$
$$= (\alpha\beta)^n \left(\prod_{i=1}^{n} t_i\right)^{\beta-1} \exp\left(-\sum_{i=1}^{n} t_i^{\beta}\right)$$

을 최대로 하는 값들이므로, α의 추정치는

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} t_i^{\beta}} \tag{21}$$

가 되나, β는 간단하게 표현되지는 않고

$$1/\beta + \sum_{i=1}^{n} t_i/n = \sum_{i=1}^{n} t_i^{\beta} \ln t_i / \sum_{i=1}^{n} t_i^{\beta} \tag{22}$$

를 만족함을 알 수 있다. 따라서 추정치를 구하기 위하여는 먼저 식(22)를 수치해석적인 방법으로 풀어 β의 값을 구한다음 이를 식(21)에 대입하여 α의 추정치를 구한다.

예 6. 예1의 데이타에 대하여 식(21), (22)에 의한 α와 β의 모수를 구해보면 $\alpha \fallingdotseq 0.0188$, $\beta \fallingdotseq 1.0$이 됨을 알 수 있다. 와이블분포의 경우에는 위와같은 방법으로 α와 β에 대한 추정치들을 구할 수 있으나 그 과

정이 번거로운 점이 없지 않다. 따라서 현실적으로는 와이블 확률지를 이용하는 편이 훨씬 용이하다. 즉 와이블 확률지에서 $Y=\ln \ln 1/(1-F(t))$ 이라하고 $X=\ln t$라고 하면, 직선식은 $Y=\ln\alpha+\beta X$의 형태이다. 따라서 β의 추정치는 타점된 점들을 잘 표현한다고 생각되는 직선을 그리고 이를 $Y=0$, $X=1$인 점(확률지에 ○표가 되어 있음)으로 옮긴 후 직선이 $X=0$일 때의 Y의 값으로 부터 읽어낼 수 있다. 한편 α의 추정치는 점들을 통과하는 직선에서 $Y=0$이면

$$\alpha=1/t_s^{\beta}$$

이 되므로 이를 이용하여 구할 수 있다. 단 t_s는 $Y=0$가 되는 t의 값이다. 예1의 데이타에서는 $t_s=58.5$이므로 α는 0.0171임을 알 수 있다.

3. 비모수적 방법

비모수적 방법은 고장시간에 대한 분포를 가정하지 않고 분석하는 방법으로서, 고장시간에 대한 가정을 하기 어려운 경우 주로 사용된다. 고장시간 데이타에 대한 비모수적 분석방법에는 누적분포에 대한 경험적 방법, 총시간타점법, 허드-존슨방법, 카플란마이어의 방법 등 여러가지가 있으나 이 절에서는 카플란-마이어 방법에 대해서만 알아 보기로 한다.

카플란-마이어 방법은 특히 관측중단된 고장시간 데이타를 분석하기 위하여 제안된 방법으로, 흔히 product-limit 추정치라고 부른다. 지금 n개의 아이템을 수명시험하여 $t_1, t_2, \cdots, t_k$와 같은 고장시간을 얻었다고 가정하자. 그리고 n_j를 t_j 직전까지 작동하였던 아이템의 수라고 하고, d_j를 t_j에서 고장난 아이템의 수라고 하면, 카플란-마이어방법은 신뢰도 함수 $R(t)$를

$$R(t)=\prod_{j:t_j<t}\frac{n_j-d_j}{n_j} \qquad (23)$$

단, $R(0)=1$

과 같이 추정하자는 것이다.

예7. 다음은 어떤 전자부품의 수명시험 데이타로 +는 관측중단된 것을 의미한다. 이를 카플란-마이어의 방법으로 신뢰도 함수의 추정치를 구하면 다음 그림6과 같이 된다.

〈Raw Data〉

31.70	39.20	57.50	65.00+	65.80
70.70	75.00+	87.50+	88.30+	94.20+
101.70+	105.80	109.20+	110.0	130.00+

@ DATA FILE NAME :ksc12.dat
@ NUMBER OF DATA :15
@ FORM OF INPUT DATA :Multiply Type I Censored Data
@ METHOD OF ALALYSIS:Nonparametric Method 〈Kaplan-Meier Method〉

〈Kaplan-Meier Calculations〉

TIME	RELIABILITY CONDITIONAL		RELIABILITY PREVIOUS		RELIABILITY CURRENT	FAILURE PROBABILITY (%)
31.70	14/15	*	1.000	=	0.933	6.667
39.20	13/14	*	0.933	=	0.867	13.333
59.50	12/13	*	0.867	=	0.800	20.000
65.00+	11					
65.80	10/11	*	0.800	=	0.727	27.273
70.70	9 /10	*	0.727	=	0.655	34.545
75.00+	8					
87.50+	7					
88.30+	6					
94.20+	5					
101.70+	4					
105.80	3 /4	*	0.655	=	0.491	50.909
109.20+	2					75.455
110.00	1 /2	*	0.491	=	0.245	
130.00+	0					

〈Reliability Plotting〉

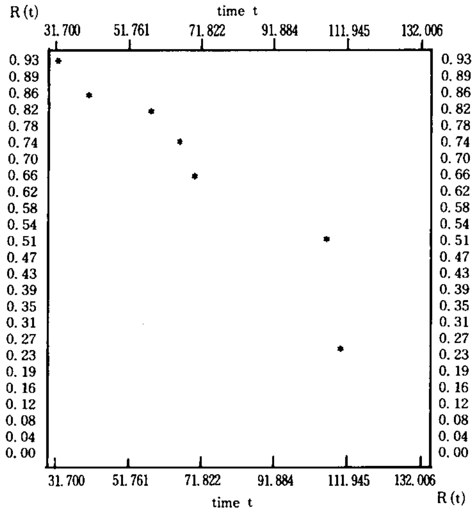


그림 6. 카플란-마이어 방법

VI. 가속수명시험 데이타의 분석

가속수명시험은 적절한 스트레스를 부가하여 고장시간을 빨리 얻을 수 있도록 하는 시험이다. 가속수명시험에 대한 설명은 앞에서도 이루어졌으므로 이 절에서는 지수분포의 고장분포함수를 갖는 경우의 데이타 분석에 관하여 알아보고자 한다.

가속수명시험을 행할 경우 데이타분석을 위하여는 보통 가속조건과 정격조건(사용조건) 하에서의 고장분포함수의 형태는 변하지 않고 다만 모수만 스트레스 수준에 따라 달라진다는 가정을 하는 것이 보통이다. 따라서 가속 조건의 스트레스 수준은 이와같은 가정이 불가능 할 경우 범위내에서만 선정되어야 하며, 이와같은 가정이 불가능한 경우에는 데이타 분석이 곤란하게 된다.

모수와 스트레스 수준간의 관계를 규정하는 가장 간단한 방법은 스트레스를 변화시켜 발생한 고장시간들 사이에 선형성을 가정하는 것이다. 즉 t_s를 스트레스 수준이 s인 가속조건에서의 고장시간이라고 하고, t_u를 가속 사용조건에서의 고장시간이라 할때

$$t_u = AF * t_s \tag{24}$$

와 같은 선형관계를 가정한 모형이 가장 일반적으로 사용되며, 여기서 AF를 가속계수 (acceleration factor)라고 부른다. (24)와 같은 가정이 성립하면 지수분포의 경우에는

$$\theta_u = AF * \theta_s \tag{25}$$

$$\lambda_u = 1/\theta_u = \lambda_s / AF$$

와 같은 관계가 성립함을 알 수 있다. 여기서 θ_u, λ_u는 사용조건에서의 MTBF와 고장률이며 θ_s와 λ_s는 가속조건에서의 값들이다. 따라서 가속계수의 값을 알고 있는 경우에는 (25)의 관계식을 이용하여 사용조건에서의 MTBF나 고장률의 값을 구할 수 있다.

예 8. 실험실에서 125℃의 온도수준에서 가속수명시험된 어떤 전자부품의 MTTE의 값을 구해보니 4500시간이었다. 정격사용조건은 25℃라 하고 두 온도 수준간의 가속계수는 35라고 한다. 그러면 이 경우 정격사용조건에서의 MTTF는 $\theta_u = 4500*35 = 157,500$ 시간 이라고 할 수 있고 따라서 고장률 $\lambda_u = 1/157,500 = 6.35*10^{-6}$(회/시간) 임을 알 수 있다.

위에서는 가속계수를 알고 있을 때의 경우를 살펴보았다. 그렇다면 이와같은 경우에는 고장시간에 대한 적절한 모형을 사용해야 하는 데 가장 많이 사용되는 것은 아레니우스 모형과 아이링 모형이다. 여기에서는 아레니우스 모형에 대하여만 알아보기로 한다. 아레니우스 모형은 θ를 MTBF라고 할 때

$$\theta = A \exp(B/kT) \tag{26}$$

와 같이 가정한 모형이다. 여기서 A와 B는 미지 모수이고 k는 볼즈만 상수, T는 절대온도를 나타낸다.

아레니우스 모형이 성립되면 가속계수는 다음과 같이 구할 수 있다. T_u를 정격사용온도, T_s를 가속조건온도라고 하면

$$AF = \theta_u/\theta_s = \exp(B/k*(1/T_u - 1/T_s)) \tag{27}$$

이 성립한다. 따라서 가속계수는 미지모수 B의 값만 데이타로 부터 추정하면 된다.

미지모수 B를 추정하기 위하여는 적어도 2수준 이상의 온도에서 수명시험을 하여 얻어진 데이타를 분석하여야 한다. 그런데 식(26)으로 부터

$$\ln\theta = \ln A + B*1/kT$$

이 성립한다. 이 식은 독립변수를 1/kT, 종속변수를 $\ln\theta$로 하는 선형회귀식을 의미한다. 따라서 A와 B의 값은 각 온도수준에서 구해진 MTBF의 값을 1/kT와 대비하여 타점된 점들을 대표하는 직선을 최소제곱추정법을 구하여 얻을 수 있다.

예 9. 85℃, 105℃, 125℃의 온도에서 수명시험을 하여 MTBF의 추정값을 각각 다음과 같이 얻었다.

온도	85℃	105℃	125℃
$\hat{\theta}$	40194	2208	242

여기서 $Y = \ln\theta$, $X = 1/kT$로 하면

온도	85℃	105℃	125℃
X	32.40	30.69	29.15
Y	10.60	7.70	5.49

가 성립한다. 따라서 선형회귀모형의 결과에 의하여

$$B = \frac{\sum X_i Y_i - n\bar{X}\bar{Y}}{\sum X_i^2 - n\bar{X}^2} = \frac{739.7865 - 3*30.75*7.93}{2841.3586 - 3*30.75^2}$$

$$= 1.58$$

이며

$$\ln A = Y - B\,X = -40.65$$
$$A = 2.22 * 10^{-18}$$

이다.

Ⅶ. 결　　언

이 글에서는 고장데이타의 수집 및 분석에 대하여 간단히 살펴보았다. 고장시간데이타의 분석 방법은 여기에서 언급한 것들 이외에도 다른 많은 방법들이 제안되어 있다. 또 고장분포함수도 지수분포와 와이블분포외의 다른 분포 즉 정규분포, 대수정규분포, 감마분포, 극치분포 등에 대한 분석법도 많이 제안되어 있다. 따라서 고장데이타분석은 각 상황에 맞는 정확한 방법을 선택하여 실시하였는데 이러한 분석의 목적은 단지 신뢰성 자료의 제시에 있는 것이 아니라 아이템의 신뢰도를 높일 수 있는 단서를 제공하는데 두어야 한다.

參 考 文 獻

[1] 한국과학기술원 응용계통실험실, 가속수명시험 설계 및 데이타분석, 미출판 보고서, 한국과학기술원 응용통계실험실, 1988.

[2] Lawless, J. F., Statistical Models and Method for Lifetime Data, Wiley, 1982.

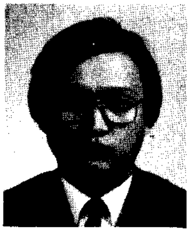
[3] Nelson, W., How to Analyze Reliability Data, ASQC, 1983.

[4] O' Connor, P. D. T. Practical Reliability Engineering, Wiley, 1988.

[5] Tobias, P. A. and Trindade, D. C., Applied Reliability, Van Nostrand Reinhold Co., 1986.

[6] US MIL-HDBK-338, Electronic Reliability Design Handbook, 1984.

筆 者 紹 介



張 重 淳
1957年 5月 27日生
1979年 2月 서울대학교 산업공학과(공학사)
1981年 2月 한국과학기술원 산업공학과(공학석사)
1986年 8月 한국과학기술원 산업공학과(공학박사)

1984年 3月～1986年 3月 아주대학교 산업공학과 전임강사
1986年 4月～1990年 3月 아주대학교 산업공학과 조교수
1987年 3月～1988年 2月 한국과학기술원 초빙교수
1990年 4月～현재 아주대학교 산업공학과 부교수