

## 임의 중단모형에서 최소제곱법을 이용한 와이블분포의 모수 추정

이우동

**요약** 임의의 기계에 대한 수명의 분포는 와이블분포를 하는 경우가 흔하다. 그리고 현실적으로 기계의 수명시간을 검정할 때, 시험시간 및 여러 환경적인 제약에 의하여 표본으로 주어진 기계의 수명을 모두 관측하기는 어렵다. 그래서, 본 연구에서는 임의 중단모형 하에서 와이블분포의 모수를 최소제곱법(least squares method)을 이용하여 추정하고 기존의 최대우도추정량(maximum likelihood estimates)과 효율성의 측면에서 비교하고자 한다.

**주제어:** 와이블분포, 최대우도함수, 최소제곱법, 생존함수

### 1. 서론

와이블 분포는 여러 수명분포 중에서 가장 많은 연구 대상이 되는 수명분포중 하나이다. 수명시간과 연관된 응용분야로 볼베어링이나 진공관, 전기인슐레이션의 분포모형이나 의학분야에서 사람에게 종양이 나타날 때까지 걸리는 시간, 쇄의 강도의 분포등 많은 분야에서 응용되는 분포이다. 특히, 와이블 분포는 누적된 충격에 의해 고장이 일어나는 기계의 수명분포에 적합하며 형상모수(shape parameter)에 따라 증가위험률(increasing failure rate), 감소위험률(decreasing failure rate), 상수위험률(constant failure rate)을 가지는 분포이다.

수명자료는 관측시 실험시간과 비용의 절약 또는 현실적인 제약조건등으로 자료를 완전히 관측하지 못하고 일부분을 중단(censoring)시킨 자료를 얻는 경우가 많다. 특히 검사 중에 있는 부품들이 우연한 사고에 의해서 관측이 중단되는 경우를 임의중단(random censoring)이라고 하고 임의중단모형하에서 와이블분포의 모수추정은 여러 학자들이 관심을 가졌고, 최대우도추정량은 표본의 크기가 증가할 때 좋은 통계적 성질을 가지고, 컴퓨터를 이용한 수치해석적 방법(Newton-Raphson method등)으로 계산 할 수 있다. Cohen, Whitten and Ding(1984)는 수정된 최대우도추정량(modified MLE)을 제안하였다. Gehan과 Siddiqui(1973)는 최소제곱법을 이용하여 지수, Gompertz, 와이블분포의 모수를 추정하였다. 최근, Lawless(1982)는 변환(log 변환)을 통하여 생존함수(survival function)와 관측된 자료를 일차직선화시킨 후, 최소제곱법을 이용하여 와이블분포와 극단치 분포(extreme value distribution)의 모수를 추정하였다.

임의 중단된 표본이 있는 경우 효율성(efficiency) 측면에서 최대우도추정량과 비슷하고 추정방법이 간단한 추정량을 찾고, 제안된 추정량과 최대우도추정량의 효율성을 모의실험(simulation)을 통하여 비교하고자 한다.

이 연구에서는 와이블분포의 생존함수(혹은 누적위험률함수)를 log변환 후 시간  $t$ 에 대해

일차함수가 되도록 만든 후, 생존함수의 비모수적 추정량과 최소제곱법을 이용하여 와이블분포의 모수를 추정한다. 2절에서는 제안된 방법을 소개하고, 3절에서는 16개의 계자권선의 수명시간을 이용하여 이 연구에서 제안된 방법으로 모수를 추정하는 적용 사례를 들고 모의실험을 통하여 제안된 방법의 효율성을 알아보고자 한다.

### 2. 와이블분포의 모수추정

형상모수  $\beta$ 와 척도모수(scale parameter)  $\alpha$ 를 가지는 와이블분포의 확률밀도함수(pdf)는

$$f(t|\alpha, \beta) = (t/\alpha)^{\beta-1} (\beta/\alpha) e^{-(t/\alpha)^\beta}, \quad 0 < t, 0 < \alpha, \beta$$

이고, 분포함수는

$$F(t) = \int_0^t f(x|\alpha, \beta) dx = 1 - \exp\{-(t/\alpha)^\beta\}, \quad t > 0$$

이다. 만약  $T$ 가 확률밀도함수가  $f(t|\alpha, \beta)$ 인 확률변수라 한다면,  $Y = \log T$ 는 다음과 같은 확률밀도함수

$$g(y|\mu, \gamma) = (1/\gamma) \exp\{(y-\mu)/\gamma\} \exp\{-\exp\{(y-\mu)/\gamma\}\}, \quad -\infty < y < \infty$$

를 가지는 극단치분포로서  $\mu = \log \alpha, \gamma = \beta^{-1}$ 이다.

$T_1, T_2, \dots, T_n$ 을 확률밀도함수가  $f(t|\alpha, \beta)$ 를 가지는 모집단으로부터 추출된  $n$ 개의 확률표본(random sample)이라 하고,  $C_1, C_2, \dots, C_n$ 을  $n$ 개의 표본에 대한 실험중단시간(censoring time)이라 두자. 그리고  $T_i$ 와  $C_i$ 는 통계적으로 독립이고  $C$ 의 분포함수를  $Q$ 라 가정하자.

임의 중단모형에서는  $Z_i = \min(T_i, C_i)$ 만을 관측할 수 있고,  $Z_i$ 의 분포함수  $H$ 는  $1-H(t) = (1-F(t))(1-Q(t))$ 인 관계가 성립한다. 그리고  $Z_i$ 의 중단지시자(censoring indicator)를

$$\delta_i = \begin{cases} 1 & T_i < C_i \\ 0 & T_i \geq C_i \end{cases}$$

라 하자. 임의중단모형에서 관찰값을  $z_1, z_2, \dots, z_n$ 이라 두고  $z_{(1)} \leq z_{(2)} \leq \dots \leq z_{(n)}$ 을  $z_1, z_2, \dots, z_n$ 의 순서통계량(order statistic)이라 하고 순서통계량  $z_{(i)}$ 에 따른 중단지시자를  $\delta_{(i)}$ 라 두자. 그리고  $\eta_i = \log c_i, y_i = \log z_i, \mu = \log \alpha, \gamma = \beta^{-1}$ 라 하자. 그러면, 임의 중단모형에서의 우도함수(likelihood function)는

$$L(\mu, \gamma) = \prod_{i=1}^n [(1/\gamma) \exp\{\frac{y_i - \mu}{\gamma}\} - \exp\{-\frac{y_i - \mu}{\gamma}\}]^{\delta_i} [\exp\{-\exp\{\frac{y_i - \mu}{\gamma}\}\}]^{1-\delta_i}$$

이다.  $r = \sum_i \delta_i$ 라 하면,  $r$ 은 중단되지 않은 자료의 개수이다.  $D = \{i|\delta_i = 1\}$  그리고  $D^c = \{i|\delta_i = 0\}$ 이라 하자. 그러면 로그-우도함수(log-likelihood function)는 아래와 같이 표현된다.

$$\log L(\mu, \gamma) = -r \log \gamma + \sum_{i \in D} (y_i - \mu)/\gamma - \sum_{i \in D} \exp\{(y_i - \mu)/\gamma\} - \sum_{i \in D^c} \exp\{(\eta_i - \mu)/\gamma\}.$$

위의 로그-우도함수를 각각  $\mu$ 와  $\gamma$ 에 대해 편미분한 뒤 0으로 두고 그 방정식을 만족하는 해가  $\mu$ 와  $\gamma$ 의 최대우도추정량  $\hat{\mu}$ 와  $\hat{\gamma}$ 이며, 와이블분포의 척도모수  $\alpha$ 와 형상모수  $\beta$ 에 대한 최

대우도추정량은 다음과 같다.

$$\hat{\alpha} = \exp\{\hat{\mu}\}, \quad \hat{\beta} = \hat{\gamma}^{-1}. \tag{1}$$

그러나 로그-우도함수(log-likelihood)를 미분한 방정식은 비선형 방정식(non-linear equation)으로 수치해석적인 방법(Newton-Raphson 방법등)으로 계산 해야 한다. 그러한 수치해석적인 방법에 의존하지 않고 최소제곱법을 이용한 추정법은 소개하면 다음과 같다. 먼저, 와이블분포의 생존함수는

$$S(t) = \exp\{-(t/\alpha)^\beta\}$$

이다. 그러므로

$$\log(-\log S(t)) = \log(\Lambda(t)) = \beta \log t - \beta \log \alpha$$

인 관계가 성립한다. 여기서  $\Lambda$ 는 확률변수  $T$ 의 누적위험률함수(cumulative hazard function)이다. 위의 식에서 모수  $\alpha$ 와  $\beta$ 를 포함하고 있는  $S(t)$ 대신에  $S(t)$ 에 대한 비모수적 추정량  $\hat{S}(t)$ 를 대입하고,

$$\varepsilon_i = \log(-\log \hat{S}(z_i)) - \log(-\log S(z_i)), i = 1, 2, \dots, n \tag{2}$$

라하면

$$\log(-\log \hat{S}(z_i)) = \beta \log z_i - \beta \log \alpha + \varepsilon_i, i = 1, 2, \dots, n$$

로 쓸 수 있다.

이제  $\hat{S}(t)$ 에 아래에 소개하는 몇 가지의 생존함수의 비모수적 추정량을 대입하여  $\alpha$ 와  $\beta$ 의 추정문제를 고려해 보자. 먼저, Kaplan과 Meier(1958)는 생존함수의 추정량(KM 추정량)을 다음과 같이 제안하였다.

$$\hat{S}_{KM}(t_0) = \prod_{i: t_{(i)} \leq t_0} \left\{ \frac{(n-i)}{(n-i+1)} \right\}^{\delta_{(i)}}. \tag{3}$$

한편, Nelson(1972)과 Aalen(1978)에 의해 제안된 누적위험률함수에 대한 경험적(empirical) 추정량으로 다음과 같다.

$$\hat{\Lambda}(t_0) = \int_0^{t_0} \frac{dN(s)}{Y(s)} = \sum_{i: t_{(i)} \leq t_0} \frac{\delta_{(i)}}{n-i+1}.$$

여기서,  $N(t_0) = \sum I\{Z_{(i)} \leq t_0, \delta_{(i)} = 1\}$ ,  $Y(t_0) = \sum I\{Z_{(i)} \geq t_0\}$ 이다. 위의 추정량을 이용한 생존함수의 추정량(NA 추정량)은 분포함수와 누적위험률함수의 관계를 이용하여 다음과 같이 제안할 수 있다.

$$\hat{S}_{NA}(t_0) = \exp\{-\hat{\Lambda}(t_0)\}. \tag{4}$$

Buckley와 James(1979)는 중단표본이 있는 경우, 아래와 같은 의사확률변수(pseudo random variable),  $T_i^*$ ,를 이용하여 생존함수의 추정량을 제안하였다.

$$T_i^* = T_i \delta_i + E[T_i | T_i \geq C_i](1 - \delta_i), i = 1, 2, \dots, n.$$

여기서  $T_i^*$ 에 포함된  $E[T_i|T_i \geq C_i]$ 의 추정량으로

$$\hat{E}[T_i|T_i \geq C_i] = C_i + \int_{C_i}^{(\infty)} \hat{S}_{KM}(s) ds / \hat{S}_{KM}(C_i)$$

를 대입하여

$$\hat{T}_i^* = T_i \delta_i + \hat{E}[T_i|T_i \geq C_i](1 - \delta_i), \quad i = 1, 2, \dots, n$$

라 두고  $\hat{T}_i^*$ 를 이용한 생존함수의 추정량은 아래와 같다.

$$\hat{S}_{BJ}(t_0) = \sum I\{\hat{T}_i^* \geq t_0\}. \tag{5}$$

생존함수의 비모수적 추정량(3),(4),(5)는 계단함수(step function)의 형태이다. 그러므로 점  $(\log Z_i, \log[-\log\{(\hat{S}(Z_i) + \hat{S}(Z_{i+1}))/2\}])$ ,  $i = 1, 2, \dots, n-1$  를 가장 근접하게 지나가는 직선을 최소제곱법으로 구한 후, 그 직선의 절편과 기울기를  $b_0$ 와  $b_1$ 이라 두면 (2)로부터 모수  $\alpha$ 와  $\beta$ 의 추정량은 아래와 같다.

$$\hat{\alpha}_l = \exp(b_0/b_1), \quad \hat{\beta}_l = b_1, \quad l = KM, NA, BJ. \tag{6}$$

여기서,  $\hat{\alpha}_{KM}$ 과  $\hat{\beta}_{KM}$ 은 KM 추정량에 기초한 최소제곱법(KM 최소제곱법)으로 구한 모수의 추정량(KM최소제곱추정량)이다.

### 3. 적용사례 및 모의실험

여기에서는 2절에서 제안된 추정방법을 이용하여 Nelson(1982)에 수록된 16개의 계자권선(field windings)의 수명시간을 관측한 자료로 최대우도 추정량과 최소제곱법을 이용한 추정량을 구해 보자. 그리고 몬테-칼로 모의실험(Monte-Carlo Simulation)을 이용하여 추정량들의 효율성을 비교해 보자.

아래의 자료는 16개의 계자권선의 수명시간을 관측한 자료이다. 이 자료에서 수명시간끝에 "+"로 표시된 자료는 중단된 자료임을 나타낸다.

31.7, 39.2, 57.5, 65.0+, 65.8, 70.0, 75.0+, 75.0+, 87.5+, 88.3+, 94.2+, 101.7+, 105.8, 109.2+, 110.0, 130.0+

이 자료를 이용하여 최대우도 추정량 (1)과 (3),(4),(5)를 대입하여 구한 추정량을 구해 보자. 추정량을 구하기 전에 자료에 대한 생존함수의 추정량(3),(4),(5)를 아래의 그림 1에 나타내었다. 그리고 16개의 자료들을 (2)로 변환 시킨 후 그 결과를 그림2에 나타내었다.

그림2로부터 log로 변환된 자료는 (2)의 직선경향이 있다는 것을 알 수 있다. 그러므로 계자권선의 수명분포가 와이블분포를 따른다고 가정해도 별 무리가 없음을 알 수 있다. 이 자료에 대한 최대우도추정량은  $\hat{\alpha} = 123.164$ ,  $\hat{\beta} = 2.376$ 이며, KM최소제곱추정량은  $\hat{\alpha}_{KM} = 134.548$ ,  $\hat{\beta}_{KM} = 1.81$  이었고, NA최소제곱추정량은  $\hat{\alpha}_{NA} = 141.428$ ,  $\hat{\beta}_{NA} = 1.754$ 이었으며, BJ최소제곱추정량은  $\hat{\alpha}_{BJ} = 103.982$ ,  $\hat{\beta}_{BJ} = 2.365$ 이었다.

다음으로 모의실험을 통하여 추정량들의 효율성을 평균제곱오차(mean squared error)측면에서 비교해 보자.

수명분포가 와이블분포인 경우, 중단분포는 지수분포로 가정하였으며 와이블분포의 형상모수를 조절하여 증가위험율( $\beta = 1.5$ ), 상수위험율( $\beta = 1.0$ ), 감소위험율( $\beta = 0.9$ )인 경우를 고려 하

였으며, 중단비율( $p$ )은 10%, 30%, 50%로 변화 시켰다. 표본의 수( $n$ )는 25, 30, 50, 100으로 변화시키며, 반복 회수를 500번으로 추정된 편기(bias)와 평균제곱오차를 계산하였다. 그리고 모의실험에 사용된 난수는 IMSL의 난수발생 부프로그램(subroutine)인 "GGUBS"를 이용하였다.

표 1은 모의실험의 계획을 요약한 것이며, 표1의 계획에 따라 모의실험을 한 결과, 표 2를 얻었다. 표 2에서 편이나 평균제곱오차가 가장 우수한 값에는 밑줄로 표시하였다.

표 2로부터 다음과 같은 결론을 얻었다.

최대우도추정량의 편이나 평균제곱오차는 이 연구에서 제안된 최소제곱추정법을 이용한 추정량들 보다는 우수하였다. 그러나 BJ 최소제곱추정량의 편이나 평균제곱오차는 중단비율과 표본수, 모수의 변화에 관계없이 KM이나 NA 최소제곱추정량보다 우수하였다.

$\alpha = 0.9, \beta = 0.9, p = 0.1, 0.5, n = 100$ 인 경우와  $\alpha = 1.0, \beta = 1.0, p = 0.1, n = 100$ 인 경우, 척도모수에 대한 BJ 최소제곱추정량의 편이나 평균제곱오차는 가장 우수하였다. 표 2를 볼 때, 표본의 크기가 증가하고 중단비율이 높아 질수록 BJ 최소제곱추정량은 편이나 평균제곱오차의 측면에서 우수한 경우가 많아진다는 사실을 알 수 있었다.

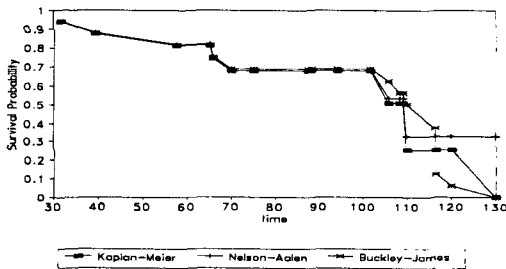


그림 1. 계자관선의 생존함수 추정량

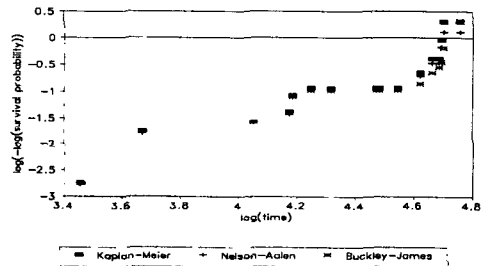


그림 2. 변환된 자료의 산포도

표 1. 모의실험계획

수명분포	중단분포	중단비율
Weibull(0.9, 0.9)	Exp(0.110)	10%
	Exp(0.456)	
	Exp(1.111)	
Weibull(1.0, 1.0)	Exp(0.111)	30%
	Exp(0.429)	
	Exp(1.000)	
Weibull(2.0, 1.5)	Exp(0.058)	50%
	Exp(0.214)	
	Exp(0.455)	

표 2. 모의실험결과

$p$	$n$		MLE	K-M	N-A	B-J	
0.1	25	Bias	.01914	-.08130	-.01580	-.10016	
		Mse	.05485	.06160	.06607	.05775	
	30	Bias	.05091	-.09424	-.11360	.09908	
		Mse	.03640	.06371	.06552	.04764	
	50	Bias	.00273	-.08688	-.03434	-.09511	
		Mse	.03859	.04359	.04224	.04322	
	100	Bias	.04112	-.07227	-.08880	-.08683	
		Mse	.02574	.05304	.05430	.03627	
	$\alpha = 0.9$ $\beta = 0.9$	25	Bias	.00937	-.04324	-.00994	-.04636
			Mse	.02652	.02921	.02991	.02706
		30	Bias	.02124	-.05780	-.06858	-.06762
			Mse	.01270	.03288	.03359	.02334
50		Bias	.00911	-.00753	.01010	-.00477	
		Mse	.01580	.01433	.01490	.01417	
100		Bias	.01320	-.05363	-.05942	-.05368	
		Mse	.00889	.01769	.01808	.01715	
0.3		25	Bias	.01482	-.05397	.02505	-.12064
			Mse	.06709	.08395	.10568	.06620
		30	Bias	.07072	-.00168	-.02452	.04459
			Mse	.04385	.15827	.15641	.04608
	50	Bias	.01450	-.03184	.03674	-.09232	
		Mse	.05056	.07453	.09612	.05295	
	100	Bias	.05328	-.01776	-.03738	-.04804	
		Mse	.03023	.12242	.12132	.04138	
	0.5	25	Bias	.00630	-.01355	.02648	-.03612
			Mse	.02873	.04173	.04897	.03027
		30	Bias	.03833	-.01582	-.02772	-.04047
			Mse	.02226	.06192	.06141	.03189
50		Bias	-.00489	-.02556	-.00643	-.00134	
		Mse	.01502	.01909	.01964	.01537	
100		Bias	.01892	.00468	-.00182	-.02178	
		Mse	.00825	.03273	.03252	.01287	
0.5		25	Bias	.03836	.05135	.19185	.20003
			Mse	.13237	.35208	.64583	.09033
		30	Bias	.08094	.13964	.11024	.03221
			Mse	.07201	.48945	.47609	.09264
	50	Bias	.05248	.05451	.16518	-.15517	
		Mse	.09619	.27812	.45421	.07059	
	100	Bias	.05469	.11420	.09001	.01301	
		Mse	.04492	.36636	.35756	.06049	
	0.5	25	Bias	.02852	.01638	.07250	-.08529
			Mse	.07910	.15567	.20550	.04266
		30	Bias	.03471	.07935	.06485	.00264
			Mse	.02337	.18727	.18409	.03207
50		Bias	.02825	.05175	.08015	.02035	
		Mse	.02621	.05796	.06812	.02140	
100		Bias	.01032	-.00869	-.01645	-.01742	
		Mse	.01545	.04985	.04949	.02237	

표 2.(계속)

$p$	$n$		MLE	K-M	N-A	B-J	
0.1	25	Bias	.01315	-.09038	-.02560	-.10781	
		Mse	.05530	.06046	.06163	.06037	
		Bias	.05763	-.09458	-.11595	-.10568	
			Mse	.04538	.08221	.08409	.05770
	30	Bias	.00071	-.08949	-.03669	-.09751	
			Mse	.03905	.04491	.04280	.04440
		Bias	.04452	-.07475	-.09299	-.09535	
			Mse	.03035	.07302	.07431	.04487
	50	Bias	.01014	-.04344	-.01005	-.04515	
			Mse	.02604	.02943	.02998	.02691
		Bias	.02362	-.05408	-.06592	-.07422	
			Mse	.01594	.05418	.05479	.02865
100	Bias	.00816	-.00865	.00903	-.00323		
		Mse	.01570	.01455	.01507	.01445	
	Bias	.01569	-.05367	-.06011	-.05842		
		Mse	.01086	.02188	.02228	.02137	
$\alpha = 1.0$ $\beta = 1.0$	25	Bias	.01445	-.05568	.02353	-.11574	
		Mse	.06605	.09148	.11537	.06724	
		Bias	.08324	.01406	-.01064	-.04539	
		Mse	.05624	.19533	.19258	.05882	
			Bias	.00917	-.04214	.02537	-.09308
				Mse	.05051	.07963	.09903
	30	Bias	.05806	.00635	-.01491	-.05378	
			Mse	.03663	.17395	.17149	.05067
		Bias	.00031	-.02640	.01263	-.03856	
			Mse	.02792	.04489	.05096	.02984
	50	Bias	.04437	.01666	.00375	-.04242	
			Mse	.02759	.10949	.10801	.03933
Bias		-.00789	-.04045	-.02168	-.00215		
		Mse	.01475	.02177	.02194	.01515	
100	Bias	.02330	.03982	.03266	-.01884		
		Mse	.01032	.05865	.05793	.01644	
	Bias	.02828	.00540	.12983	-.18989		
		Mse	.11858	.31964	.51293	.09204	
0.5	25	Bias	.08841	.23001	.19861	.02825	
		Mse	.09110	.73079	.70983	.11614	
		Bias	.04990	.01330	.11793	-.14217	
		Mse	.10582	.33053	.58983	.06957	
			Bias	.06535	.22908	.20318	.01735
				Mse	.05631	.57132	.55565
	30	Bias	.02383	-.01388	.03798	-.07320	
			Mse	.07231	.14801	.18653	.04293
		Bias	.04149	.15125	.13589	.00384	
			Mse	.03037	.29719	.29126	.04048
	50	Bias	.02491	.03617	.06388	.02754	
			Mse	.02720	.06678	.07697	.02537
Bias		.01217	.02236	.01404	-.01518		
		Mse	.01910	.08629	.08535	.02825	

표 2.(계속)

$p$	$n$		MLE	K-M	N-A	B-J	
0.1	25	Bias	.00145	-.14476	-.05911	-.16269	
		Mse	.09441	.11825	.11144	.11183	
	30	Bias	.09327	-.10344	-.13389	-.14539	
		Mse	.10220	.22888	.23104	.13283	
	50	Bias	-.00725	-.13275	-.06105	-.13968	
		Mse	.06834	.08676	.07866	.08276	
	100	Bias	.06891	-.10006	-.12575	-.13846	
		Mse	.06971	.19405	.19658	.10261	
	$\alpha = 2.0$ $\beta = 1.5$	25	Bias	.00674	-.07407	-.03001	-.06669
			Mse	.04775	.06088	.06006	.05188
		30	Bias	.03185	-.03701	-.05377	-.10809
			Mse	.03586	.18061	.18053	.06681
50		Bias	.00599	-.02029	.00321	-.01282	
		Mse	.02838	.02751	.02787	.02576	
100	Bias	.02187	-.06061	-.06980	-.08224		
	Mse	.02184	.06768	.06821	.04717		
0.3	25	Bias	.00055	-.17687	-.09009	-.15869	
		Mse	.10941	.17158	.17776	.13178	
	30	Bias	.12964	.35364	.32262	-.05599	
		Mse	.12802	1.06436	1.03625	.13498	
	50	Bias	.00088	-.15004	-.07519	-.12558	
		Mse	.09149	.16280	.17317	.10351	
	100	Bias	.09108	.30886	.28225	-.06226	
		Mse	.08375	.88539	.86548	.11380	
	0.5	25	Bias	-.00575	-.13007	-.08603	-.05574
			Mse	.05141	.09585	.09605	.05443
		30	Bias	.06721	.31322	.29628	-.04219
			Mse	.05522	.61219	.59994	.09098
50		Bias	-.01184	-.12372	-.10188	-.01715	
		Mse	.02804	.05581	.05319	.02597	
100	Bias	.03592	.25069	.24175	.00173		
	Mse	.02209	.30877	.30402	.03899		
0.5	25	Bias	.00252	-.19550	-.08984	-.22085	
		Mse	.15139	.63397	.77566	.16462	
	30	Bias	.15143	.88314	.84770	.07721	
		Mse	.20230	2.65377	2.58236	.30354	
	50	Bias	.03311	-.18726	-.09653	-.15709	
		Mse	.14254	.54742	.73306	.12125	
	100	Bias	.11395	.91633	.88684	.05277	
		Mse	.13577	2.45553	2.39463	.20078	
	0.5	25	Bias	.00143	-.22945	-.18290	-.06537
			Mse	.08514	.31941	.34643	.08069
		30	Bias	.07442	.73979	.72188	.03151
			Mse	.06961	1.44658	1.41704	.11120
50		Bias	.01498	-.22632	-.20333	.03798	
		Mse	.03950	.15538	.15363	.04030	
100	Bias	.02360	.51123	.50171	.01759		
	Mse	.04021	.70792	.69693	.07366		



#### 4. 결론

와이블분포의 모수에 대한 최대우도추정량은 편기나 평균제곱오차의 측면에서 우수한 추정량임은 이미 알려져 있다. 그러나 최대우도추정량을 얻기 위해서는 수치해석적인 비선형방정식의 해법에 의존해야만 한다.

이 연구에서 제안된 추정법은 계산하기가 쉽고 (3),(4),(5)만 계산 된다면 (2)를 이용하여 기존의 통계package(SPSS,MINITAB등)를 사용하여 추정량을 얻을 수 있다는 장점이 있다.

특히, BJ최소제곱추정량은 NA나 KM최소제곱추정량보다 편기나 평균제곱오차의 측면에서 우수하였고 중단비율이 높고 표본의 수가 많은 경우는 최대우도추정량보다 더 우수한 경우도 있었다. 그러므로 중도절단된 자료가 많은 경우, BJ추정법에 기초한 최소자승법을 이용한다면 와이블분포의 모수추정에서는 최대우도추정량보다 간단하고 효율성도 뛰어난 추정량을 얻을 수 있다고 기대된다.

#### 참고문헌

- Aalen, O.O.(1978), *Nonparametric Inference for a family of Counting Processes*, Annals of Statistics, Vol.6, pp.701-726.
- Buckley, J. and James, I.(1979), *Linear Regression with Censored Data*, Biometrika, Vol.66, pp.429-436.
- Cohen, A.C. Whitten, B.J. and Ding, Y.(1984), *Modified Moment Estimator for the Three-parameter Weibull Distribution*, Journal of Quality Technology, Vol.10, pp.159-167.
- Gehan, E.A. and Siddiqui, M.M.(1973), *Simple Regression Methods for Survival Time Studies*, Journal of the American Statistical Association, Vol.63, No.334, pp.848-856.
- Kaplan, E.L. and Meier, P. (1958), *Nonparametric Estimator from Incomplete Observations*, Journal of the American Statistical Association, Vol.53, pp.457-481.
- Lawless, J. F. (1982), *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, Inc., New York.
- Nelson, W. (1972), *Theory and Applications of Hazard Plotting for Censored Failure Data*, Technometrics, Vol.14, pp.945-965.
- Nelson, W. (1982), *Applied Life Data Analysis*, John Wiley & Sons, Inc., New York.

# An Estimation of Parameters in Weibull Distribution Using Least Squares Method under Random Censoring Model

Woo Dong Lee

**Abstract** In this paper, under random censorship model, an estimation of scale and shape parameters in Weibull lifetime model is considered. Based on nonparametric estimator of survival function, the least square method is proposed. The proposed estimation method is simple and the performance of the proposed estimator is as efficient as maximum likelihood estimators. An example is presented, using field winding data. Simulation studies are performed to compare the performances of the proposed estimator and maximum likelihood estimator.

**Keywords** : Weibull Distribution, Maximum Likelihood Estimates, Least Square Estimates, Survival Function.