



# 제 8 강. 생존분석2

이재원 교수  
고려대학교 통계학과

# 주요내용

1. 비모수적 방법을 이용한 생존함수의 비교
2. 모수적 방법을 이용한 생존함수의 추정과 비교
3. 준모수적 방법을 이용한 생존함수의 추정과 비교



# 비모수적 방법을 이용한 생존함수의 비교

# 비모수적 방법을 이용한 생존함수의 비교 <예 6.3>

흑색종(melanoma) 환자들에 대한 BCG 와 CP(coryne-bacterium parvum)의 생존지속 효과를 비교하기 위한 연구에서 30명의 흑색종 환자 중 11명은 BCG 처리를 받고 나머지 19명은 CP처리를 받았다고 한다.

자료는 <표 6.4>과 같다. 중도절단이 포함되어 있는 경우에 이 두 그룹의 생존분포를 비교하기 위한 방법을 알아보자. (p. 244)

# 비모수적 방법을 이용한 생존함수의 비교

그룹1  $T_{11}, T_{12}, \dots, T_{1n_1} \stackrel{iid}{\sim} S_1, C_{11}, C_{12}, \dots, C_{1n_1} \stackrel{iid}{\sim} G_1$

그룹2  $T_{21}, T_{22}, \dots, T_{2n_2} \stackrel{iid}{\sim} S_2, C_{21}, C_{22}, \dots, C_{2n_2} \stackrel{iid}{\sim} G_2$

그룹1  $\Rightarrow (x_{11}, \delta_{11}), (x_{12}, \delta_{12}), \dots, (x_{1n_1}, \delta_{1n_1})$

그룹2  $\Rightarrow (x_{21}, \delta_{21}), (x_{22}, \delta_{22}), \dots, (x_{2n_2}, \delta_{2n_2})$

$x_{ij} = \min(T_{ij}, C_{ij}), \delta_{ij} = 0$  이면 중도절단

# 정해진 시점에서의 생존률의 비교

가 설

$$H_0 : S_1(t_k) = S_2(t_k)$$

$$H_1 : S_1(t_k) \neq S_2(t_k)$$

검정통계량

$$Z = \frac{\hat{S}_1(t_k) - \hat{S}_2(t_k)}{\sqrt{\hat{Var}(\hat{S}_1(t_k)) + \hat{Var}(\hat{S}_2(t_k))}} \sim N(0, 1) \quad (6.8)$$

# Log-rank test (로그-순위 검정)

가 설

$$H_0 : S_1 = S_2 \quad H_1 : S_1 \neq S_2$$

$$t_1 < t_2 < \dots < t_k$$

	사망	생존	계
그룹 1	$d_{1i}$	$n_{1i} - d_{1i}$	$n_{1i}$
그룹 2	$d_{2i}$	$n_{2i} - d_{2i}$	$n_{2i}$
계	$d_i$	$n_i - d_i$	$n_i$

→  $n_{1i}, n_{2i}, d_i$  고정 ,  $d_{1i} \sim \text{hypergeometric}$

# Log-rank test (로그-순위 검정)

$$E(d_{1i}) = E_{1i} = n_{1i} d_i / n_i$$

$$Var(d_{1i}) = V_{1i} = \frac{n_{1i} n_{2i}}{n_i - 1} \frac{d_i (n_i - d_i)}{n_i^2} \quad (6.9)$$

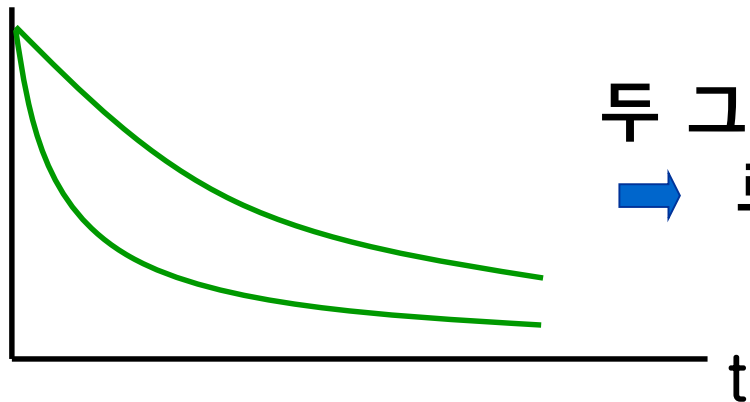
$$T = \frac{\left\{ \sum_{i=1}^k (d_{1i} - E_{1i}) \right\}^2}{\sum_{i=1}^k V_{1i}} \sim \chi^2(1) \quad \text{under } H_0 \quad (6.10)$$

$$T = \frac{\left\{ \sum_{i=1}^k w_i (d_{1i} - E_{1i}) \right\}^2}{\sum_{i=1}^k w_i^2 V_{1i}} \sim \chi^2(1) \quad \text{under } H_0 \quad (6.11)$$



# Log-rank test (로그-순위 검정)

생존 확률



두 그룹의 위험률의 비가 상수  
→ 로그-순위 검정법 선호(가중치=1)

두 그룹의 위험률의 비가 상수가 아닌 경우  
→ 다른 가중치를 주는 방법 선호(Gehan, Prentice 등)

예) Gehan 검정통계량의 경우  $w_i = \frac{n_i}{N+1}$

# R Program

```
library(survival)
흑색종환자자료=read.table("f:WWWORKWW흑색종환자자료.txt",header=T)
흑색종환자자료
attach(흑색종환자자료)

# 누적한계추정값(Kaplan-Meier추정값)
적합2=survfit(Surv(time,status)~x,data=흑색종환자자료)
summary(적합2)
```

- **survfit 함수 (survival 패키지)** : ~ 좌변에 Surv(시간, 절단상태)를 입력하고 ~ 우변에는 처리효과, 공변량을 입력.

# R Program (계속)

```
# 로그-순위 검정법(log-rank test)
```

```
survdif(Surv(time, status)~x, data=흑색종환자자료)
```

```
# Gehan-Wilcoxon 검정
```

```
library(npsm)
```

```
gehan.test(time, status, x)
```

```
# 누적한계추정값 비교 그래프
```

```
plot(적합2,xlab="시간",ylab="생존함수",lty=c(1,2),mark.time=T)
```

```
legend(5,0.2,c("CP 처리 그룹","BCG 처리 그룹"),lty=c(2,1))
```

- **survdif 함수 (survival 패키지)** : 두 개 이상의 생존곡선을 검정하는 함수. ~ 좌변에 Surv(시간, 절단상태)를 입력하고 ~ 우변에는 처리효과, 공변량을 입력.
- **gehan.test 함수 (npsm 패키지)** : Gehan-Wilcoxon 검정을 실시하는 함수. 생존시간, 절단상태, 공변량 순으로 입력.

# R 출력결과

```
> 적합2=survfit(Surv(time,status)~x,data=흑색종환자자료)
> summary(적합2)
Call: survfit(formula = Surv(time, status) ~ x, data = 흑색종환자자료)
```

x=BCG

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
3.9	11	1	0.909	0.0867	0.754	1.000
5.4	10	1	0.818	0.1163	0.619	1.000
7.9	9	1	0.727	0.1343	0.506	1.000
10.5	8	1	0.636	0.1450	0.407	0.995
19.5	4	1	0.477	0.1755	0.232	0.981

x=CP

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
6.9	19	1	0.947	0.0512	0.852	1.000
7.7	18	1	0.895	0.0704	0.767	1.000
8.0	16	1	0.839	0.0854	0.687	1.000
8.3	13	1	0.774	0.1003	0.601	0.998
24.4	3	1	0.516	0.2211	0.223	1.000

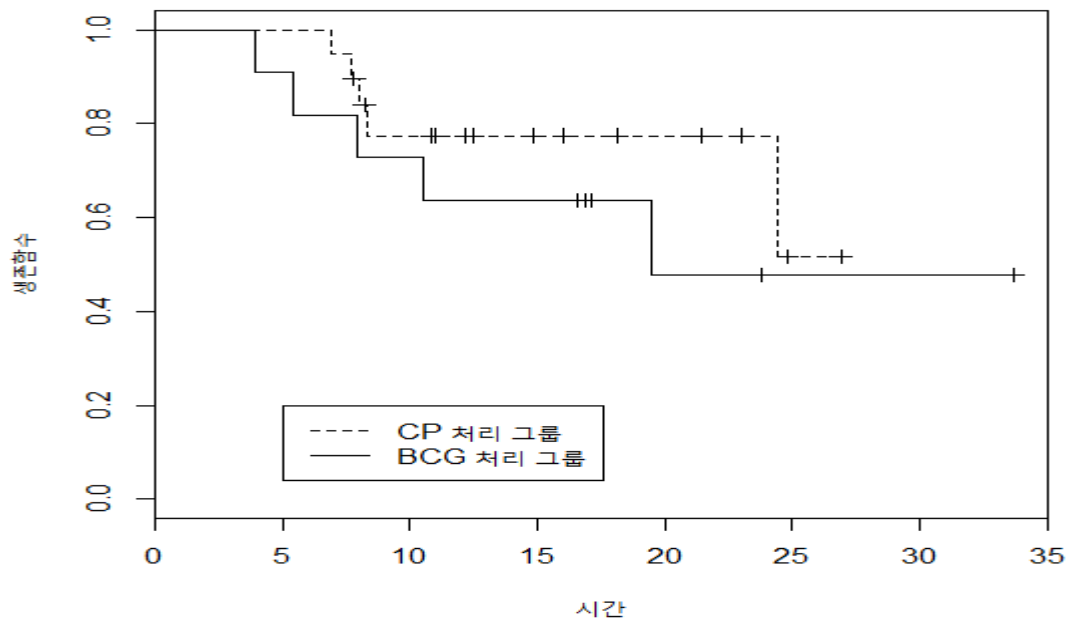
# R 출력결과 (계속)

```
> library(npsm)
필요한 패키지를 로딩중입니다: Rfit
> # 로그-순위 검정법(log-rank test)
> survdiff(Surv(time, status)~x,data=흑색종환자자료)
Call:
survdiff(formula = Surv(time, status) ~ x, data = 흑색종환자자료)

      N Observed Expected (O-E)^2/E (O-E)^2/V
x=BCG  11         5      3.68     0.469     0.747
x=CP   19         5      6.32     0.274     0.747

Chisq= 0.7  on 1 degrees of freedom, p= 0.387
> library(npsm)
> # Gehan-Wilcoxon 검정
> gehan.test(time,status,x)
statistic = -0.9917953 , p-value = 0.3212974
```

# 흑색종 환자들의 추정된 생존함수 그래프 <그림 6.5>





# 모수적 방법을 이용한 생존함수의 추정과 비교

# 모수적 모형 (1)

## ■ 지수분포(exponential distribution)

- 위험률이 시간에 대해 일정한 경우.

$$h(t) = \lambda [\text{또는 } \log h(t) = \mu \text{ 형태}]$$

$$\Rightarrow S(t) = e^{-\lambda t}, \quad f(t) = \lambda e^{-\lambda t}$$

## ■ 고펜르츠 분포(Gompertz distribution)

- **로그위험함수**를 시간에 대한 선형함수로 가정.

교재 251페이지에 “위험함수”를  
“로그위험함수”로 수정.

$$\Rightarrow \log h(t) = \mu + \alpha t$$



# 모수적 모형 (2)

## ■ 와이블 분포(Weibull distribution)

$$\log h(t) = \mu + \alpha \log t \text{ [또는 } h(t) = \lambda t^\alpha \text{ 형태]}$$

▶ 공변량  $x_1, \dots, x_k$  가 있을 때

$$\text{지수분포 : } \log h(t) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k,$$

$$\text{곰페르츠 분포 : } \log h(t) = \alpha t + \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k, \quad (6.12)$$

$$\text{와이블 분포 : } \log h(t) = \alpha \log t + \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k$$

## <예 6.4>

여기서 사용된 자료는 Prentice(1973)에 소개된 것으로 40명의 폐암 환자의 생존시간을 조사한 것이다. 40명의 환자 중 21명은 기존 치료방법인 처리 1에, 나머지 19명은 새로운 치료방법인 처리 2에 할당되었으며, 생존시간에 영향을 미칠 것으로 생각되는 공변량은 다음과 같다.

### 공변량

$x_1$  : 진단시의 환자상태(Performance Status:0~100점)

$x_2$  : 환자의 나이(단위:년)

$x_3$  : 진단 후 연구 참여시까지의 시간 (단위:월)

TYPE : 종양의 유형(squamous, small, adeno, large)

# R Program

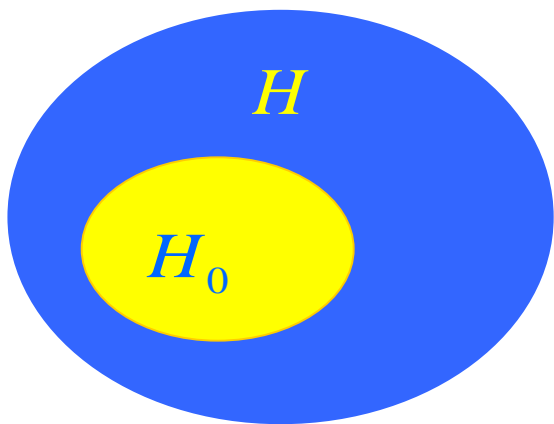
```
폐암환자자료=read.table("f:WORKWW폐암환자자료.txt",header=T)
head(폐암환자자료)
```

```
# 모수적 모형에 근거한 생존시간 분석
```

```
와이블=survreg(Surv(time,status)~x1+x2+x3+factor(trt)+factor(type),data=폐암환자자료,dist="weibull")
summary(와이블)
```

- **survreg 함수 (survival 패키지)** : 모수적 생존모형을 적합하는 함수. ~ 좌변에 Surv(시간, 절단상태)를 입력하고 ~ 우변에는 처리효과와 공변량을 입력. dist에는 생존시간이 따르는 분포를 입력. "weibull", "exponential", "gaussian", "logistic", "lognormal", "loglogistic" 등을 입력할 수 있음.

# 우도비검정 (Likelihood Ratio Test)



$$-2\log \frac{L(H_0)}{L(H)} > \chi_{\alpha}^2(d.f.) \text{ 이면 } H_0 \text{ 기각}$$

$d.f.$  = 추정하는 모수의 갯수의 차

(예)  $H_0 : \beta_1 = \cdots = \beta_p = 0$

# 폐암 환자의 모수적 모형적합을 위한 R출력결과 (일부)

```
Weibull distribution  
Loglik(model)= -203.4    Loglik(intercept only)= -219.7  
    Chisq= 32.56 on 7 degrees of freedom, p= 3.2e-05  
Number of Newton-Raphson Iterations: 6  
n= 40
```

→ Log Likelihood for WEIBULL=-203.4

→  $2(\log L - \log L_0) = 32.56 > 14.067 = \chi^2_{0.95}(7)$

→ 모든 공변량의 회귀계수가 0 이라는 귀무가설 기각

# 폐암 환자의 모수적 모형적합을 위한 R 출력결과 (일부)

```
> 와이بل=survreg(Surv(time,status)~x1+x2+x3+factor(trt)+factor(type),  
+ data=폐암환자자료,dist="weibull")  
> summary(와이بل)
```

Call:

```
survreg(formula = Surv(time, status) ~ x1 + x2 + x3 + factor(trt) +  
factor(type), data = 폐암환자자료, dist = "weibull")
```

	Value	Std. Error	z	p
(Intercept)	1.21253	1.26642	0.957	3.38e-01
x1	0.05420	0.00961	5.642	1.68e-08
x2	0.00920	0.01758	0.523	6.01e-01
x3	0.00412	0.01040	0.396	6.92e-01
factor(trt)2	0.26547	0.34633	0.767	4.43e-01
factor(type)2	-0.50676	0.45069	-1.124	2.61e-01
factor(type)3	-1.25710	0.49716	-2.529	1.15e-02
factor(type)4	-0.38117	0.39469	-0.966	3.34e-01
Log(scale)	-0.13656	0.13165	-1.037	3.00e-01

Scale= 0.872

# R 출력결과 설명

▶ 모형식 :  $\log T = 1.213 + 0.054x_1 + 0.009x_2 + 0.004x_3 +$  (6.13)  
 $+ 0.265x_4 - 0.507x_5 - 1.257x_6 - 0.381x_7 + 0.872e$

▶  $x_5, x_6, x_7$  : 종양의 유형을 가리키는 가변수

$x_5 = 1$  if 'small',  $= 0$  o.w.

$x_6 = 1$  if 'adeno',  $= 0$  o.w.

$x_7 = 1$  if 'large',  $= 0$  o.w.

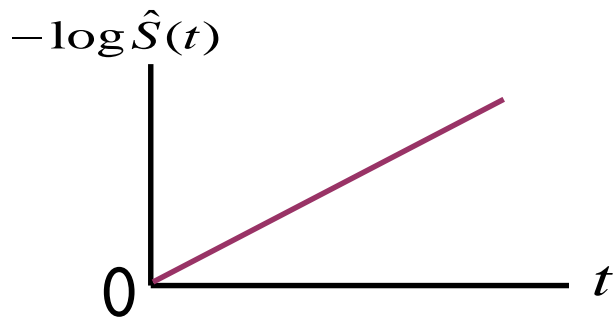
▶  $x_4$  : 처리그룹을 나타내는 가변수(dummy variable)  
처리가 standard일 때 1, test일 때 0의 값을 가짐.

# 모수적 모형의 적합도 검토

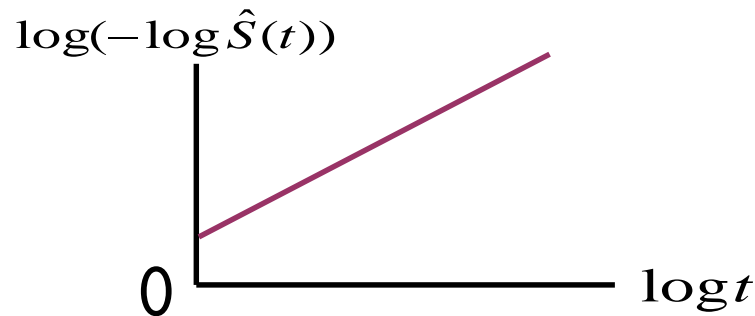
-각 모형에서의 AIC(Akaike Information Criteria) 비교

AIC 값이 가장 작은 모형

-그래프 이용



$T \sim \text{Exponential}$



$T \sim \text{Weibull}$



# 와이블 분포와 일반화감마 분포 비교를 위한 R Program

```
library(flexsurv)
일반화감마분포=flexsurvreg(formula=Surv(time,status)~x1+x2+x3+factor(trt)+
factor(type),data=폐암환자자료,dist="gengamma")
일반화감마분포
와이블분포= flexsurvreg(formula=Surv(time,status)~x1+x2+x3+factor(trt)+
factor(type) ,data=폐암환자자료,dist="weibull")
와이블분포

plot(와이블분포,xlab="시간",ylab="생존함수")
lines(일반화감마분포,col="blue",lty=2)
legend("topright",c("와이블분포","일반화감마분포"),lty=1:2,col=c("red","blue"))
```

- **flexsurv 패키지** : Flexible parametric survival models.
- **flexsurvreg 함수 (flexsurv 패키지)** : 확장된 분포 하에서의 생존모형 적합.

survreg함수와 동일한 문법을 가짐. ~ 좌변에는 Surv(시간, 절단상태)와 같이 입력하고, 우변에는 공변량 및 처리효과를 입력, dist에는 생존시간이 따르는 분포를 입력

# 와이블 분포와 일반화 감마분포를 비교하기 위한 R 출력결과

```
> 일반화감마분포=flexsurvreg(formula=Surv(time,status)~x1+x2+x3+factor(trt)+  
+ factor(type), data=폐암환자자료, dist="gengamma")  
> 일반화감마분포
```

Call:

```
flexsurvreg(formula = Surv(time, status) ~ x1 + x2 + x3 + factor(trt) +  
factor(type), data = 폐암환자자료, dist = "gengamma")
```

Estimates:

	data	mean	est	L95%	U95%	se	exp(est)
mu		NA	1.34298	-1.19193	3.87789	1.29335	NA
sigma		NA	0.80446	0.46458	1.39300	0.22535	NA
Q		NA	1.26231	-0.28258	2.80720	0.78822	NA
x1	56.50000		0.05426	0.03591	0.07261	0.00936	1.05576
x2	56.57500		0.00986	-0.02409	0.04381	0.01732	1.00991
x3	15.65000		0.00554	-0.01614	0.02722	0.01106	1.00556
factor(trt)2	0.47500		0.24953	-0.43745	0.93651	0.35051	1.28342
factor(type)2	0.27500		-0.60115	-1.61112	0.40882	0.51530	0.54818
factor(type)3	0.12500		-1.39383	-2.56794	-0.21973	0.59904	0.24812
factor(type)4	0.25000		-0.49078	-1.44164	0.46007	0.48514	0.61215

N = 40, Events: 37, Censored: 3

Total time at risk: 5784

Log-likelihood = -203.3927, df = 10

AIC = 426.7853

# 와이블 분포와 일반화 감마분포를 비교하기 위한 R 출력결과 (계속)

```
> 와이블분포=flexsurvreg(formula=Surv(time,status)~x1+x2+x3+factor(trt)+  
+ factor(type),data=폐암환자자료,dist="weibull")  
> 와이블분포
```

Call:

```
flexsurvreg(formula = Surv(time, status) ~ x1 + x2 + x3 + factor(trt) +  
factor(type), data = 폐암환자자료, dist = "weibull")
```

Estimates:

	data mean	est	L95%	U95%	se	exp(est)
shape	NA	1.14632	0.88556	1.48386	0.15095	NA
scale	NA	3.36198	0.28142	40.16345	4.25475	NA
x1	56.50000	0.05420	0.03539	0.07301	0.00960	1.05569
x2	56.57500	0.00920	-0.02523	0.04363	0.01757	1.00924
x3	15.65000	0.00412	-0.01625	0.02449	0.01039	1.00413
factor(trt)2	0.47500	0.26547	-0.41332	0.94426	0.34633	1.30404
factor(type)2	0.27500	-0.50676	-1.39000	0.37649	0.45064	0.60245
factor(type)3	0.12500	-1.25710	-2.23142	-0.28279	0.49711	0.28448
factor(type)4	0.25000	-0.38117	-1.15468	0.39233	0.39465	0.68306

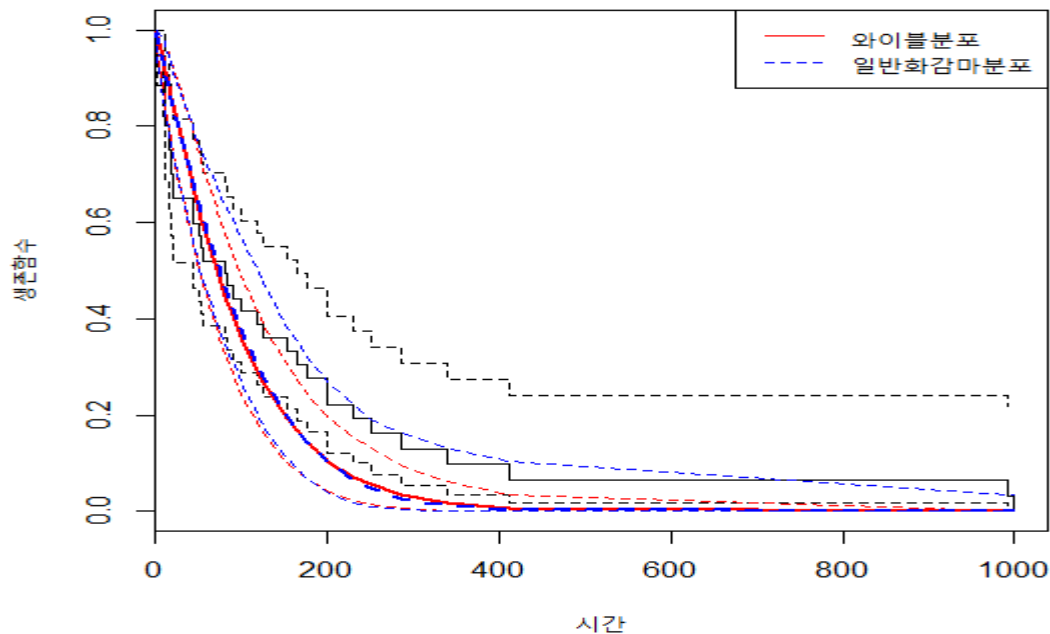
N = 40, Events: 37, Censored: 3

Total time at risk: 5784

Log-likelihood = -203.449, df = 9

AIC = 424.898

# 와이블 분포와 일반화 감마분포를 비교하기 위한 생존함수 그래프 <그림 6.6>



# 와이블 모형의 적합성 검토를 위한 R Program



이 한 줄을 교재 258페이지에  
추가해야 합니다

```
폐암누적한계추정=survfit(Surv(time,status)~1,data=폐암환자자료)
```

```
생존함수추정=폐암누적한계추정$surv
```

```
시간 = 폐암누적한계추정$time
```

```
LT = log(시간)
```

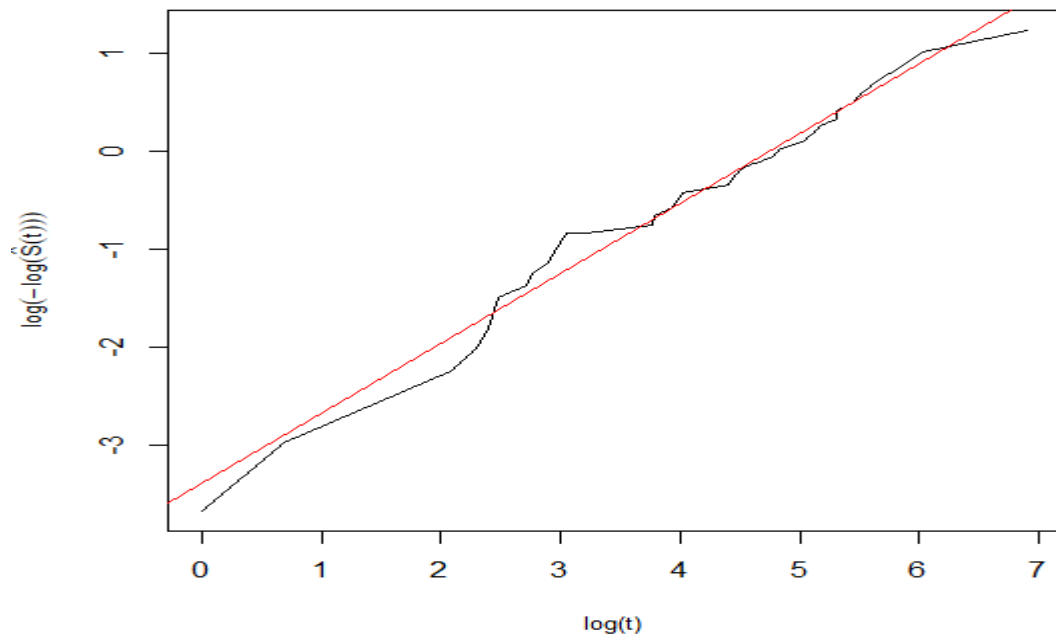
```
LL = log(-log(생존함수추정))
```

```
plot(LT, LL, type = "l", xlab = "log(t)", ylab = expression(log(-log(hat(S)(t)))),
```

```
cex.lab = 0.8)
```

```
abline(lm(LL[-37] ~ LT[-37])$coef, col="red")
```

# 와이블 모형의 적합성 검토를 위한 그래프 <그림 6.7>





# 준모수적 방법을 이용한 생존함수의 추정과 비교

# Cox의 비례위험모형

$$h(t) = h_0(t) \exp(\beta_1 x_1 + \cdots + \beta_k x_k) \quad (6.14)$$

여기서  $h(t)$  : 공변량들을 고려했을 때의 위험함수

$h_0(t)$  : 기저위험함수

$$\log h(t) = \alpha(t) + \beta_1 x_1 + \cdots + \beta_k x_k \quad (6.15)$$

$$h_i(t) / h_j(t) = \exp\{\beta_1 (x_{i1} - x_{j1}) + \cdots + \beta_k (x_{ik} - x_{jk})\} \quad (6.16)$$



# 폐암자료의 비례위험모형에 대한 분석을 위한 R program

# Cox비례위험모형에 근거한 생존시간 분석

```
coxfit1=coxph(Surv(time,status)~x1+x2+x3+factor(trt)+factor(type),data=폐암환  
자자료, ties="breslow")  
summary(coxfit1)
```

# 생존함수 추정값

```
적합4=survfit(coxfit1)  
summary(적합4)
```

# 생존함수 그래프

```
plot(적합4)
```

# 누적위험함수 그래프

```
plot(적합4, fun="cumhaz")
```

# R 프로그램 설명

- **coxph 함수 (survival 패키지)** : Cox의 비례위험 모형을 적합함.  
~ 좌변에 Surv(시간, 절단상태)를 입력하고 ~ 우변에는 처리효과와 공변량을 입력.
- **survfit 함수 (survival 패키지)** : Cox 비례위험 모형에서 생존함수를 추정함.  
fit4에서 4번째 열(fit4\$surv)은 생존함수의 추정치를 나타내는데, 여기에  $-\log()$ 를 취해 주면 누적위험함수가 됨.
- **plot 함수 (survival 패키지)** : 인자로 survfit 함수의 출력 결과를 입력하면, survfit 함수를 통하여 추정한 생존함수의 그래프를 그림. fun = "" 이 지정되지 않으면 추정된 생존함수의 그래프를, fun = "cumhaz"으로 지정된 경우 추정된 누적위험함수의 그래프를 그려줌.

# 폐암자료의 비례위험모형에 대한 분석을 위한 R 출력결과

```
> coxfit1=coxph(Surv(time,status)~x1+x2+x3+factor(trt)+factor(type),
+ data=폐암환자자료, ties="breslow")
> summary(coxfit1)
Call:
coxph(formula = Surv(time, status) ~ x1 + x2 + x3 + factor(trt) +
      factor(type), data = 폐암환자자료, ties = "breslow")

n= 40, number of events= 37

              coef exp(coef)    se(coef)      z Pr(>|z|)
x1          -0.059717  0.942031  0.013834 -4.317 1.58e-05 ***
x2          -0.011513  0.988553  0.020707 -0.556  0.5782
x3           0.000543  1.000543  0.011806  0.046  0.9633
factor(trt)2 -0.398939  0.671031  0.406388 -0.982  0.3263
factor(type)2  0.307582  1.360132  0.543616  0.566  0.5715
factor(type)3  1.378104  3.967373  0.626980  2.198  0.0279 *
factor(type)4  0.332843  1.394929  0.484384  0.687  0.4920
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
x1              0.9420      1.0615    0.9168    0.9679
x2              0.9886      1.0116    0.9492    1.0295
x3              1.0005      0.9995    0.9777    1.0240
factor(trt)2    0.6710      1.4902    0.3026    1.4882
factor(type)2   1.3601      0.7352    0.4687    3.9474
factor(type)3   3.9674      0.2521    1.1610   13.5578
factor(type)4   1.3949      0.7169    0.5398    3.6046

Concordance= 0.765  (se = 0.058 )
Rsquare= 0.518    (max possible= 0.994 )
Likelihood ratio test= 29.2  on 7 df,  p=0.0001328
Wald test            = 25.73  on 7 df,  p=0.0005628
Score (logrank) test= 30.12  on 7 df,  p=9.028e-05
```

# 폐암자료의 비례위험모형에 대한 분석을 위한 R 출력결과

```
> # 생존할수 추정값
> 적합4=survfit(coxfit1)
> summary(적합4)
Call: survfit(formula = coxfit1)

   time  n.risk  n.event  survival   std.err  lower 95% CI  upper 95% CI
1      1      40        1  9.89e-01  1.13e-02   9.67e-01  1.000
2      2      39        1  9.78e-01  1.71e-02   9.45e-01  1.000
8      8      38        2  9.56e-01  2.66e-02   9.05e-01  1.000
10     10      36        1  9.38e-01  3.28e-02   8.76e-01  1.000
11     11      35        1  9.20e-01  3.85e-02   8.48e-01  0.999
12     12      34        2  8.85e-01  4.78e-02   7.96e-01  0.984
15     15      32        1  8.65e-01  5.28e-02   7.68e-01  0.975
16     16      31        1  8.46e-01  5.74e-02   7.40e-01  0.966
18     18      30        1  8.23e-01  6.23e-02   7.10e-01  0.955
19     19      29        1  7.98e-01  6.73e-02   6.76e-01  0.941
20     20      28        1  7.69e-01  7.23e-02   6.39e-01  0.924
21     21      27        1  7.36e-01  7.69e-02   6.00e-01  0.903
43     43      25        1  7.00e-01  8.17e-02   5.56e-01  0.879
44     44      24        1  6.63e-01  8.57e-02   5.15e-01  0.854
51     51      23        1  6.28e-01  8.89e-02   4.76e-01  0.829
54     54      22        1  5.87e-01  9.18e-02   4.32e-01  0.797
56     56      21        1  5.47e-01  9.35e-02   3.91e-01  0.765
82     82      20        1  5.09e-01  9.42e-02   3.54e-01  0.732
84     84      19        1  4.69e-01  9.51e-02   3.15e-01  0.697
90     90      18        1  4.30e-01  9.50e-02   2.78e-01  0.663
100    100      17        1  3.86e-01  9.46e-02   2.39e-01  0.624
118    118      15        1  3.40e-01  9.40e-02   1.98e-01  0.584
126    126      14        1  2.97e-01  9.18e-02   1.62e-01  0.544
153    153      13        1  2.57e-01  8.85e-02   1.30e-01  0.505
164    164      12        1  2.17e-01  8.44e-02   1.01e-01  0.465
177    177      11        1  1.81e-01  7.88e-02   7.73e-02  0.425
200    200      10        1  1.41e-01  7.14e-02   5.25e-02  0.381
201    201       9        1  1.08e-01  6.24e-02   3.45e-02  0.336
```

...(생략)

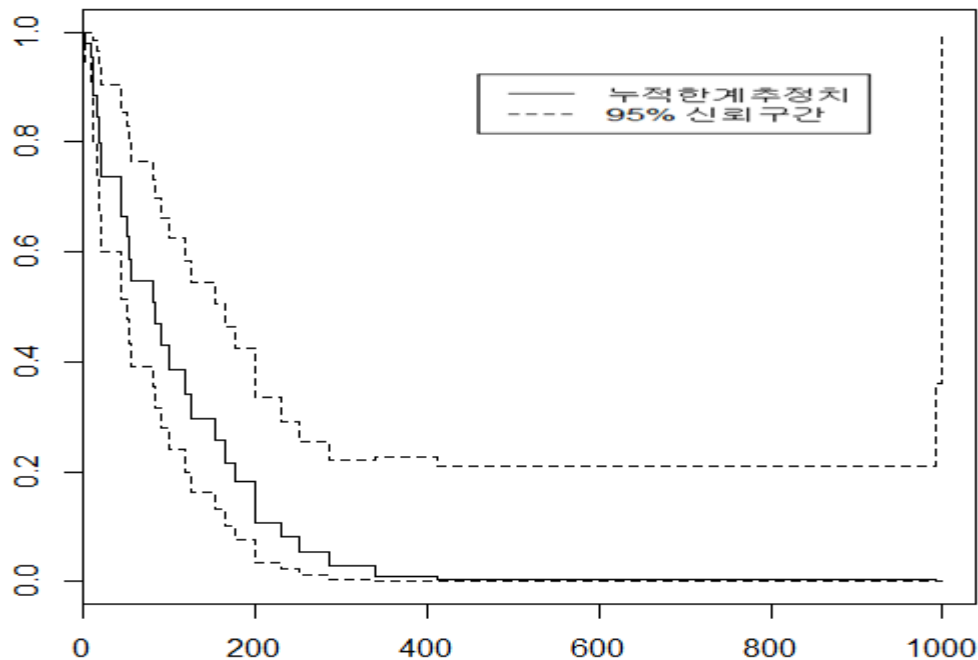
# 비례위험모형에서의 생존함수 추정

$$S(t) = [S_0(t)]^{\exp(\beta_1 x_1 + \dots + \beta_k x_k)} \quad (6.17)$$

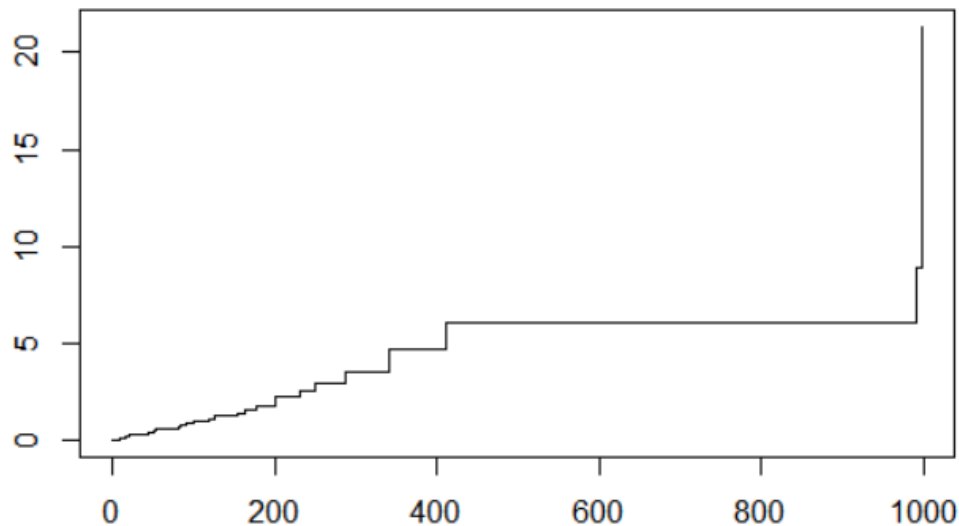
$S(t)$  : 공변량들이 있을때의 생존함수

$S_0(t)$  : 공변량이 모두 0일 때의 기저생존함수

# 폐암자료의 모형에 대한 분석에서 추정된 생존함수 그래프<그림 6.8>



# 폐암자료의 비례위험모형에 대한 분석 에서 추정된 누적위험함수 그래프 <그림 6.9>



# 비례성 검토를 위한 R program

# 비례성 검토를 위한 로그-로그 그림

```
coxfit2=coxph(Surv(time,status)~x1+x2+x3+strata(trt)+factor(type),data=폐암환자자료)
survfit(coxfit2)
처리=폐암환자자료$trt
생존함수추정2=survfit(coxfit2)$surv
시간2=survfit(coxfit2)$time
LT=log(시간2)
LL=log(-log(생존함수추정2))
plot(LT[1:19], LL[1:19], type = "s", xlab = "log(t)",
      ylab = expression(log(-log(hat(S)(t)))), cex.lab = 0.8, lwd = 2,
      xlim = c(0,7), ylim = c(-4,3), col = "red")
lines(LT[20:37], LL[20:37], type = "s", xlab = "log(t)", ylab = expression(log(-
log(hat(S)(t)))), cex.lab = 0.8, lwd = 2)
legend(1,2.5,c("처리 1","처리 2"), col = c("red", "black"), lty = 1)
```

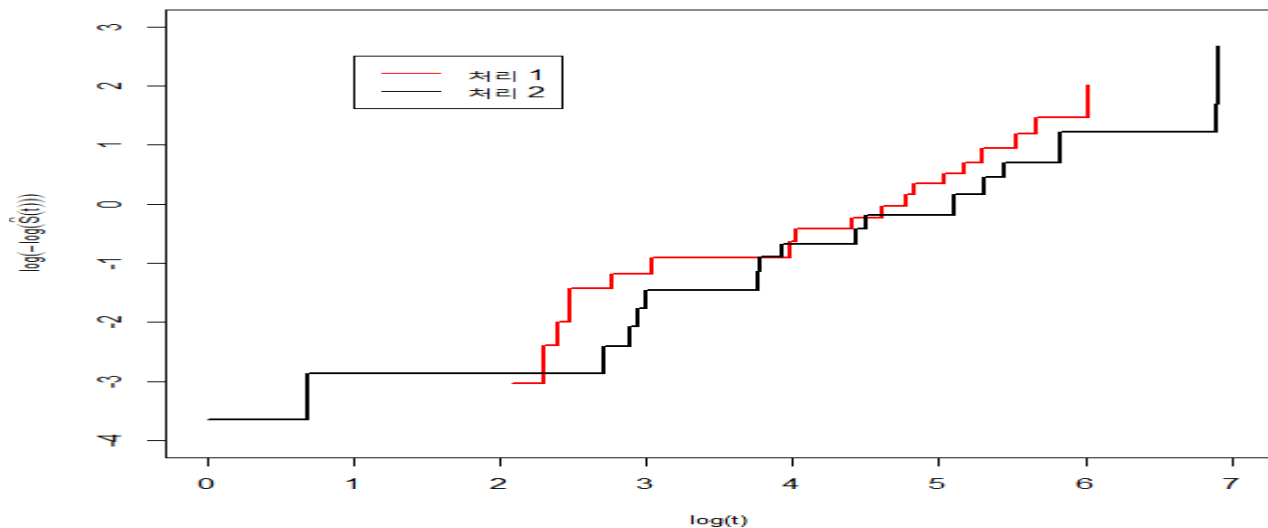
교재 코드에 오차가 있습니다.  
빨간색으로 된 글자를 주의하세요.

폐암환자 처리별로 나누어 그래프를 그림

- **coxph 함수 (survival 패키지)** : Cox의 비례위험 모형을 적합하는 함수.  
~ 우변에서 처리효과를 strata(trt)과 같이 입력해 주면, 처리를 층으로 입력하여,  
처리1과 처리2로 나누어 각각 Cox 비례위험 모형을 적합함.



# 비례성 검토를 위한 로그-로그 그림 <그림 6.10>



# 핵심정리

1. 비모수적 방법을 이용한 생존함수의 비교
  - 로그순위검정
2. 모수적 방법을 이용한 생존함수의 추정과 비교
  - 모수모형, 적합도검정
3. 준모수적 방법을 이용한 생존함수의 추정과 비교
  - Cox 비례위험모형, 비례성 검토



## 제8강

수고하셨습니다!

