

7-5.

(2)

다음과 같은 SAS 코드를 통해 당뇨병 환자의 모수적 모형을 적합시킨다. 분포는 와이블 분포를 가정하기로 한다.

```
DATA diabetes;
INPUT censor time age bmi dage smoke @@;
CARDS;
1 3.6 63 25.1 46 1 1 15.4 71 26 59 0 1 11.3 51 32 49 1
1 10.3 59 28.1 57 1 1 5.8 50 26.1 49 1 0 8 66 45.3 49 0
1 14.6 42 30 41 1 1 11.4 40 35.7 36 1 1 7.2 67 28.1 61 0
1 5.5 86 32.9 61 0 1 11.1 52 37.6 46 1 1 16.5 42 43.4 37 0
1 10.9 60 25.4 60 0 1 2.5 75 49.7 57 1 0 10.8 81 35.2 81 0
1 4.7 60 37.3 39 0 0 5.5 60 26 42 0 1 4.5 63 21.8 60 1
1 9 62 18.2 43 0 1 6.8 57 34.1 41 1 0 3.6 71 25.6 54 1
1 12.1 58 35.1 45 0 1 8.1 42 32.5 28 1 1 11.1 45 44.1 40 0
0 7 66 29.7 59 1 1 1.5 61 29.2 54 0 1 11.7 48 25.2 30 1
1 0.3 82 25.3 50 0
;

PROC LIFEREG DATA=diabetes;
CLASS smoke;
MODEL time*censor(0) = age bmi dage smoke/ DIST=WEIBULL;
RUN;
```

분석 결과는 다음과 같다.

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	3.2683	0.6987	1.8989	4.6377	21.88	<.0001
age	1	-0.0410	0.0141	-0.0686	-0.0135	8.50	0.0035
bmi	1	-0.0044	0.0138	-0.0315	0.0227	0.10	0.7497
dage	1	0.0282	0.0131	0.0025	0.0539	4.61	0.0317
smoke	0 1	0.3697	0.2170	-0.0556	0.7949	2.90	0.0884
smoke	1 0	0.0000	.	.	.	.	.
Scale	1	0.4185	0.0729	0.2975	0.5887		
Weibull Shape	1	2.3893	0.4159	1.6986	3.3609		

생존 함수 모형식은

$$\log T = 3.27 - 0.041 * age - 0.0044 * bmi + 0.0282 * dage + 0.37 * smoke$$

으로 표현할 수 있다. 여기서 흡연경력이 없을 때는  $smoke = 0$ 이고 없을 때는  $smoke = 1$ 이다. 회귀계수  $\beta$ 를 해석하면 진단 시 나이가 1만큼 낮을수록  $\log T$ 는 0.0282만큼 커지고, 흡연경력이 없을 경우  $\log T$ 는 0.3697만큼 커진다. 반대로 현재 나이가 1만큼 높을수록  $\log T$ 는

0.041만큼 낮아지고 BMI가 1만큼 낮을수록  $\log T$ 는 0.0044만큼 낮아진다. 이 때 나이(age)와 진단 시 나이(dage)에 대한 회귀 계수는 p-value가 각각 0.0035와 0.0317로 5% 유의수준 하에서 유의하다. 그러나 BMI(bmi)와 흡연 여부(smoke)에 대한 회귀 계수는 p-value가 각각 0.7497과 0.0884로 생존시간에 유의한 영향을 미치지 않는다.

7-6.

다음과 같은 SAS 코드를 통해 비례위험모형을 이용하여 자료를 분석하기로 한다.

**DATA** acute;

INPUT time censor age bone @@;

cards;

18 1 0 0 9 1 0 1 28 0 0 0 31 1 0 1

39 0 0 1 19 0 0 1 45 0 0 1 6 1 0 1

8 1 0 1 15 1 0 1 23 1 0 0 28 0 0 0

7 1 0 1 12 1 1 0 9 1 1 0 8 1 1 0

2 1 1 1 26 0 1 0 10 1 1 1 4 1 1 0

3 1 1 0 4 1 1 0 18 1 1 1 8 1 1 1

3 1 1 1 14 1 1 1 3 1 1 0 13 1 1 1

13 1 1 1 35 0 1 0

;

**run**;

**PROC PHREG** DATA=acute;

MODEL time\*censor(0) =age bone ;

**RUN**;

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	131.656	126.230
AIC	131.656	130.230
SBC	131.656	132.501

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	5.4260	2	0.0663
Score	5.4212	2	0.0665
Wald	5.0979	2	0.0782

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
age	1	1.01317	0.45740	4.9065	0.0268	2.754
bone	1	0.35025	0.43917	0.6360	0.4252	1.419

회귀계수에 대한 출력 결과를 보면, 나이(age)에 대한 회귀계수가 1.01317로 주어지고, p값이 유의수준 0.05보다 낮으므로 유의하다. 위험비(Hazard Ratio)를 보면 나이가 50세 이상일 경우 다른 조건이 동일한 환자와 비교할 때 2.754(=exp(1.01317))배 위험하다는 것을 알 수 있

다. 반면 골수 응고 100%의 여부는 그 p값이 0.4252로, 환자의 생존시간에 유의한 영향을 미치지 않는다.