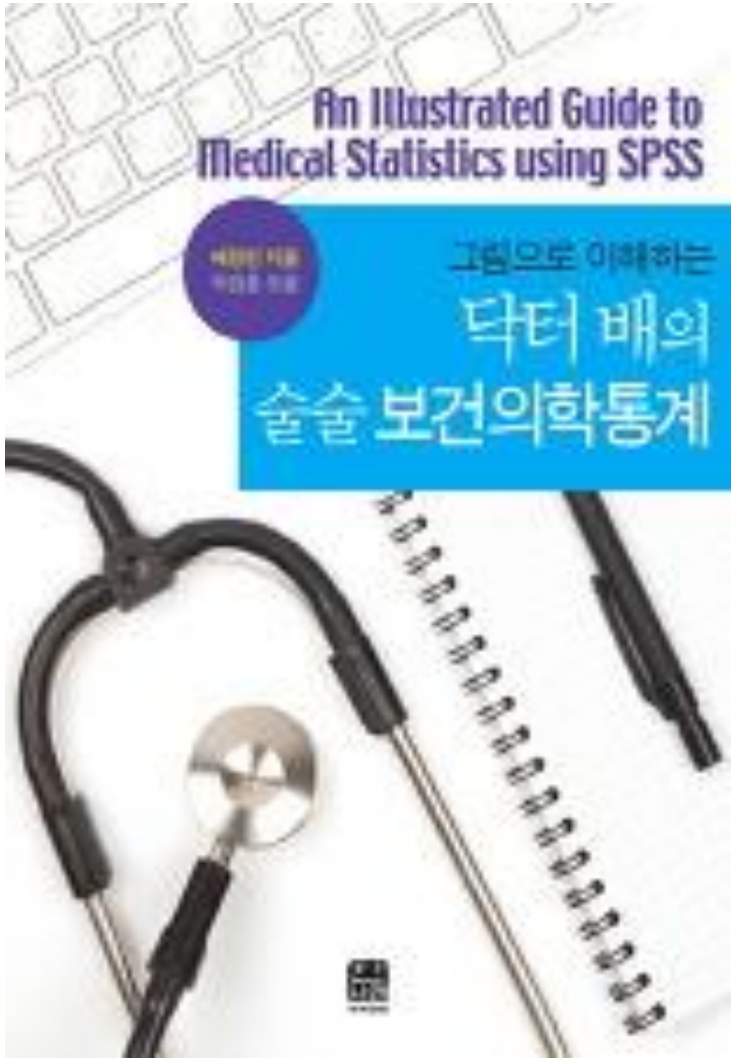
(그림으로 이해하는)
닥터 배의
술술 보건의학통계

라 가 영



07.
생존율의 추정 및 군의 생존율 비교

- 가. 생존 연구의 준비
- 나. Kaplan-Meier 생존분석
- 다. 로그순위법

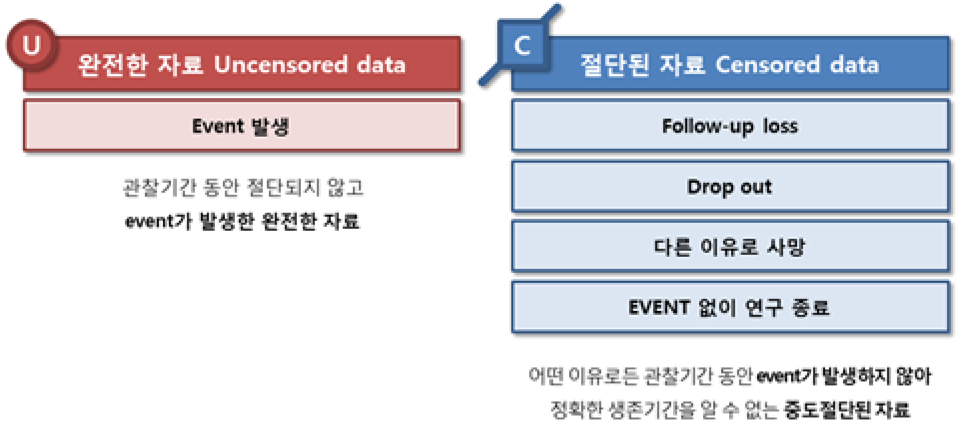
07. 생존율의 추정 및 군의 생존율 비교

생존분석

- **생존분석**(survival analysis) : 시간에 따른 사망이나 재발 등의 변화를 관찰하는 분석
- **사건**(event) : 생존분석에서 '사망'이나 '재발'과 같이 연구자가 관심을 갖고 있는 변화
- **사건 발생까지의 시간**(time to event) : 생존분석에서 분석의 대상
- **Kaplan-Meier 생존분석**(Kaplan-Meier curve analysis) : 특정 집단의 생존율을 분석
- **로그순위법**(Log-rank test) : 두 집단의 생존율을 비교
- **Cox의 비례위험모형**(Cox's proportional hazard model) : 생존율에 영향을 미치는 위험인자를 분석(11장에서 공부)

07.
생존율의 추정 및 군의 생존율 비교

가. 생존 연구의 준비

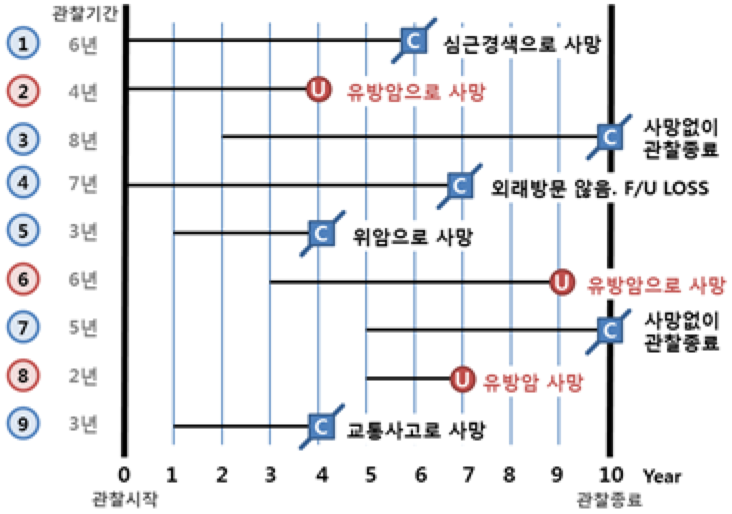


- **완전한 자료**(uncensored data) : 생존자료의 관찰기간 동안 사건(event)이 발생한 자료를 사건 발생까지 절단되지 않았을 경우

- **절단된 자료**(censored data) : 사건이 발생되기 전에 자료 수집이 종료된 모든 경우를 사건 발생 전에 관찰한 경우

07. 생존율의 추정 및 군의 생존율 비교

가. 생존 연구의 준비



- **절단된 자료**들은 사건 발생까지 걸린 시간은 알 수 없지만, 최소한 이 관찰 기간 동안 사건이 발생하지 않았다는 정보를 제공
- 개체의 총 관찰 시간을 **사건 발생까지 걸린 시간**으로 표현하고 그 기간 동안 사건이 발생하지 않았음을 표기하여 자료를 정리

07.
생존율의 추정 및 군의 생존율 비교

가. 생존 연구의 준비

- 생존분석에서 자료를 코딩하는 방법

- 생존분석 자료에서 관찰의 시작과 종료 시점은 중요하지 않음.
- **사건 발생 여부**(censored of uncensored)와 **사건 발생까지 걸린 시간**(time to event)만 분석에 이용
- 다음 표는 난소암 환자 26명을 2군으로 나누어 표준 치료법을 시행한 군(n = 13)과 새 치료법을 적용한 군(n = 13)의 생존율을 비교한 자료

07. 생존율의 추정 및 군의 생존율 비교

가. 생존 연구의 준비

- 생존분석에서 자료를 코딩하는 방법

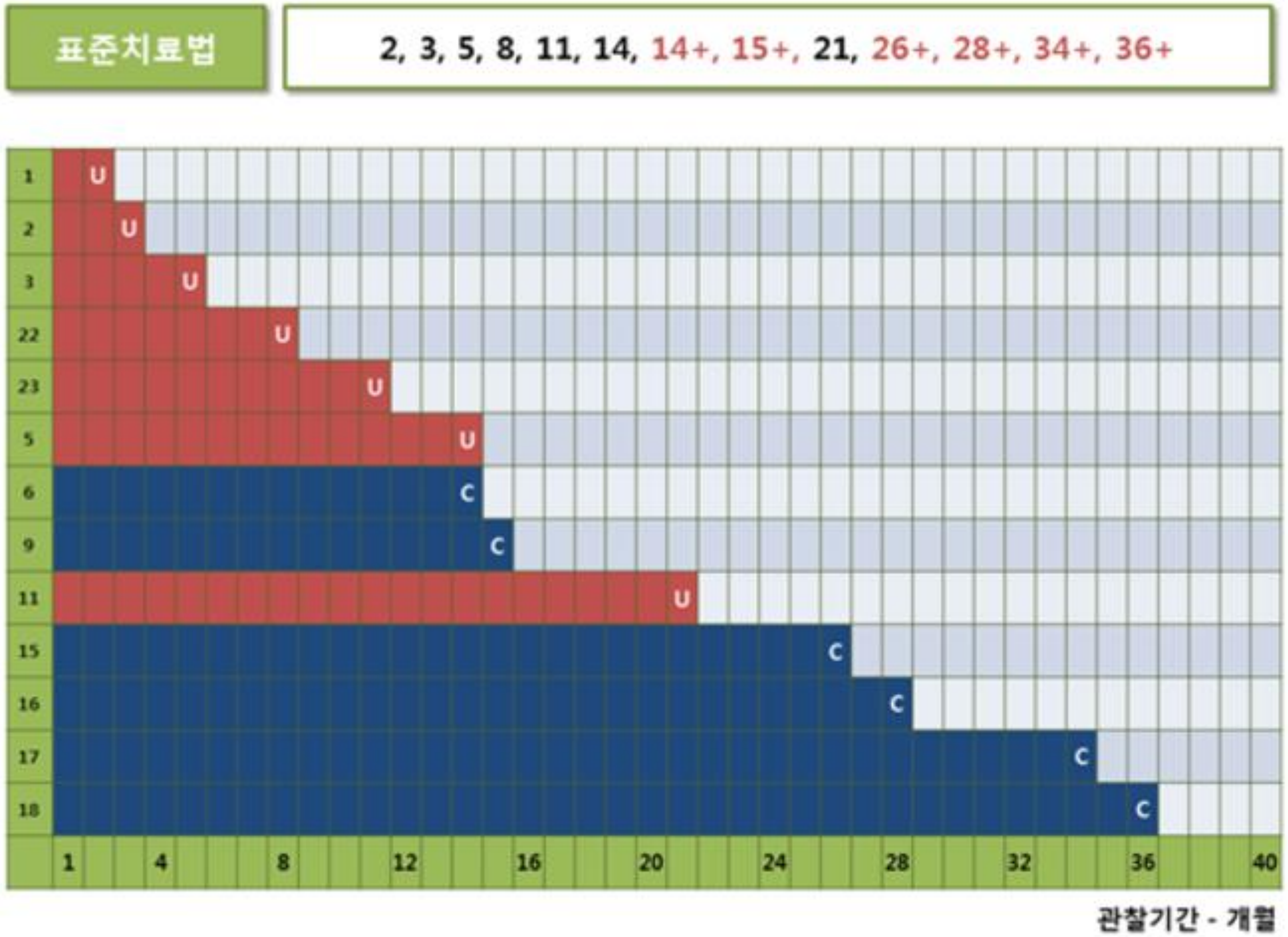
	1, 표준치료법		0, Censored
	2, 새 치료법	Time to event	1, Death

증례번호	치료군	관찰기간(개월)	사망
1	1	2	1
2	1	3	1
3	1	5	1
4	2	14	0
5	1	14	1
6	1	14	0
7	2	15	1

07.
생존율의 추정 및 군의 생존율 비교

가. 생존 연구의 준비

- 생존분석에서 자료를 코딩하는 방법

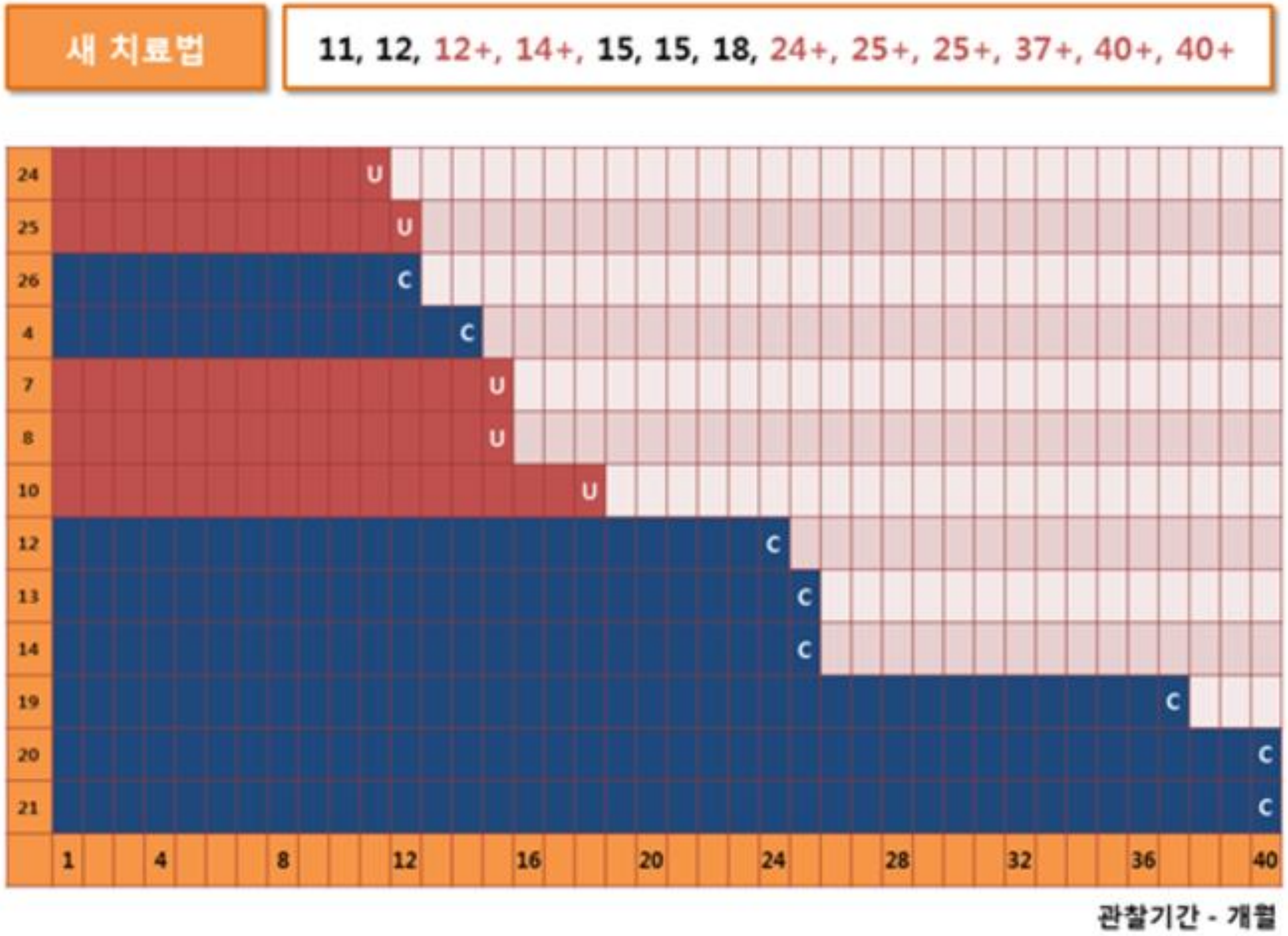


관찰 기간이 작은 순서부터 정렬
절단된 자료는 아직 사망이 발생한 것이 아니므로 기간에 +를 덧붙여서 완전한 자료와 구분

07.
생존율의 추정 및 군의 생존율 비교

가. 생존 연구의 준비

- 생존분석에서 자료를 코딩하는 방법



관찰 기간이 작은 순서부터 정렬
절단된 자료는 아직 사망이 발생한 것이 아니므로 기간에 +를 덧붙여서 완전한 자료와 구분

07.
생존율의 추정 및 군의 생존율 비교

나. Kaplan-Meier 생존분석

- **생명표법**(life table method) : 일정한 간격마다 구간생존율을 구하는 방법. 표본의 크기가 50이 넘어야 적용 가능

- **Kaplan-Meier 생존분석**(Kaplan-Meier curve analysis) : 사건(사망)이 발생한 시점마다 구간생존율을 구하여 이들의 누적으로서 누적 생존율을 구하는 방식. 표본의 크기가 작아도 적용 가능. 누적한계추정법(product-limit method)라고도 불림.

07.
생존율의 추정 및 군의 생존율 비교

나. Kaplan-Meier 생존분석

- 사건(사망)이 발생한 시점마다 생존율을 계산
- 관찰 기간의 순서대로 자료를 정리한 뒤 각 구간별로 관찰대상 수 중 생존자 수의 분율로 구간생존율(P(t))를 계산
- 누적생존율(S(t))은 각 구간별 구간생존율을 차례로 곱함으로써 추정

- Kaplan-Meier 생존곡선을 통해 특정 시점의 생존율(1년, 2년 생존율)을 추정할 수 있으며, 반대로 생존율이 50%로 떨어지는 시점도 추정 가능.

$$P(t) = \frac{t\ \text{시점의 생존자수}}{t\ \text{시점의 관찰대상수}}$$

$$S(t) = S(t-1) \times P(t)$$

07.

생존율의 추정 및 군의 생존율 비교

나. Kaplan-Meier 생존분석

표준치료법

증례번호	관찰기간 (월)	이 시점의 생존자수	이 시점의 관찰대상수	구간생존율 P(t)	누적생존율 S(t)
1	2	12	13	$^{12}/_{13} = 0.923$	0.923
2	3	11	12	$^{11}/_{12} = 0.917$	0.846
3	5	10	11	$^{10}/_{11} = 0.909$	0.769
22	8	9	10	$^{9}/_{10} = 0.900$	0.692
23	11	8	9	$^{8}/_{9} = 0.889$	0.615
5	14	7	8	$^{7}/_{8} = 0.875$	0.538
6	14+	7	7	1.0	·
9	15+	6	6	1.0	·
11	21	4	5	$^{4}/_{5} = 0.800$	0.431
15	26+	4	4	1.0	·
16	28+	3	3	1.0	·
17	34+	2	2	1.0	·
18	36+	1	1	1.0	·

절단된 증례의 경우에는 사건이 발생하지 않았으므로 구간 생존율은 1.0

07.
생존율의 추정 및 군의 생존율 비교

나. Kaplan-Meier 생존분석

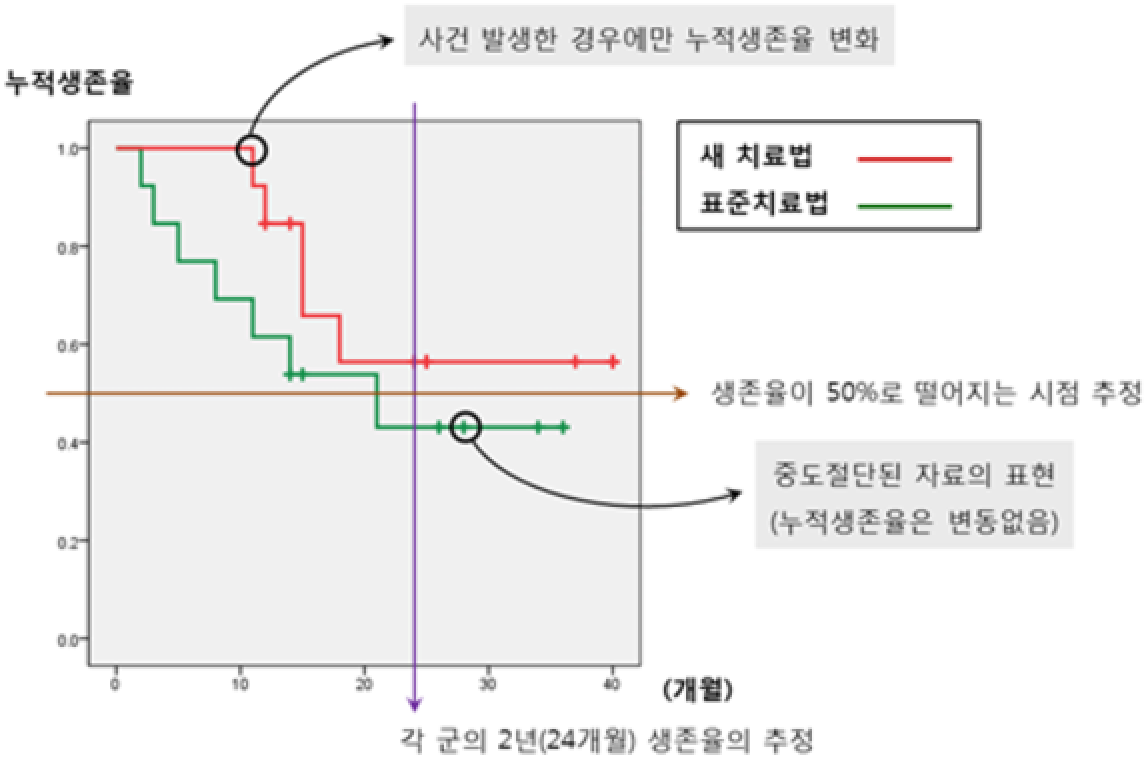
새 치료법

증례번호	관찰기간 (월)	이 시점의 생존자수	이 시점의 관찰대상수	구간생존율 P(t)	누적생존율 S(t)
24	11	12	13	$^{12}/_{13} = 0.923$	0.923
25	12	11	12	$^{11}/_{12} = 0.917$	0.846
26	12+	11	11	1.0	"
4	14+	10	10	1.0	"
7	15	8	9	$^{8}/_{9} = 0.889$	0.752
8	15	7	8	$^{7}/_{8} = 0.875$	0.658
10	18	6	7	$^{6}/_{7} = 0.857$	0.563
12	24+	6	6	1.0	"
13	25+	5	5	1.0	"
14	25+	4	4	1.0	"
19	37+	3	3	1.0	"
20	40+	2	2	1.0	"
21	40+	1	1	1.0	"

절단된 증례의 경우에는 사건이 발생하지 않았으므로 구간 생존율은 1.0

07. 생존율의 추정 및 군의 생존율 비교

나. Kaplan-Meier 생존분석



07.
생존율의 추정 및 군의 생존율 비교

다. 로그순위법

- 두 군을 합한 전체 집단을 관찰 기간순으로 배열하고 이중 절단된 항목을 모두 지우고 사건(사망)이 발생한 구간만 남겨서 정리
- 각 구간의 사망에 대한 기대빈도를 계산

07. 생존율의 추정 및 군의 생존율 비교

다. 로그순위법

$$1\times\frac{13}{13+13}=0.5$$

$$1\times\frac{13}{13+13}=0.5$$

관찰기간 (개월)	관찰대상수		표준치료법 사망수		새 치료법 사망수	
	표준치료법	새 치료법	관찰빈도	기대빈도	관찰빈도	기대빈도
2	13	13	1	0.500	-	0.500
3	12	13	1	0.480	-	0.520
5	11	13	1	0.458	-	0.542
8	10	13	1	0.435	-	0.565
11	9	13	1	0.818	1	1.182
12	8	12	-	0.400	1	0.600
14	7	10	1	0.412	-	0.588
15	6	9	-	0.800	2	1.200
18	5	7	-	0.417	1	0.583
21	5	6	1	0.455	-	0.545
계			**7**	**5.175**	**5**	**6.825**

$$2\times\frac{6}{6+9}=0.8$$

$$2\times\frac{9}{6+9}=1.2$$

07. 생존율의 추정 및 군의 생존율 비교

다. 로그순위법

- 전체 관찰빈도와 기대빈도의 관계는 자유도가 1인 x^2 분포를 따름
- 검정통계량 x^2값이 3.84보다 큰 경우 귀무가설을 기각하고 두 군은 유의한 차이를 보임
- 예) x^2 값이 1.132 < 3.84이므로 두 군의 생존곡선이 차이가 있다고 말 할 수 없음

$$\chi^2 = \sum \frac{(\text{관찰사망수 합} - \text{기대사망수 합})^2}{\text{기대사망수 합}} = \frac{(7-5.175)^2}{5.175} + \frac{(5-6.825)^2}{6.825} = 1.132 < 3.84$$

$$\rightarrow p > 0.05$$