

여러 집단에 대한 로그-순위 검정

여러 집단에 대한 로그-순위 검정

$K(\geq 2)$ 개 모집단에 대한 생존함수의 동일성 검정

관심 가설

$$H_0 : h_1(t) = h_2(t) = \cdots = h_K(t), \quad t \leq \tau$$

H_1 : 적어도 $t \leq \tau$ 의 일부분에서 한 개의 $h_g(t)$ 는 다르다.

데이터 구조

$K(\geq 2)$ 개 모집단에서 발생한 데이터

- $t_1 < t_2 < \cdots < t_D$ 개의 구별된 관측값을 가짐 (D 개 위험집합)
- 시점 t_i 에 g 번째 모집단에서 d_{gi} 개 사건 발생하고 위험집합에 Y_{gi} 개체가 있음
- $d_i = \sum_{g=1}^K d_{gi}$: t_i 시점에 관측된 사건발생수
- $Y_i = \sum_{g=1}^K Y_{gi}$: t_i 시점에서 통합한 데이터에서 위험개체수
- $\frac{d_i}{Y_i}$: t_i 시점에서 관측된 위험률
- $\frac{d_{gi}}{Y_{gi}}$: t_i 시점에서 그룹 g 에서 관측된 위험률

귀무가설이 참이라면 $\frac{d_i}{Y_i} = \frac{d_{1i}}{Y_{1i}} = \cdots = \frac{d_{Ki}}{Y_{Ki}}$

여러 집단에 대한 로그-순위 검정

$$\sum_{g=1}^K (\hat{h}_g - \bar{h}) = \sum_{g=1}^K \left\{ \left(\sum_{i=1}^D W_g(t_i) \frac{d_{gi}}{Y_{gi}} \right) - \left(\sum_{i=1}^D W_g(t_i) \frac{d_i}{Y_i} \right) \right\}$$

귀무가설 하에서 t_i 시점에서 g 번째 모집단의 위험률 추정량 $\frac{d_{gi}}{Y_{gi}}$ 과 모집단 전체에서 통합된 위험률 추정량 $\frac{d_i}{Y_i}$ 의 차이에 기반한 검정통계량

$$Z_g(\tau) = \sum_{i=1}^D W_g(t_i) \left(\frac{d_{gi}}{Y_{gi}} - \frac{d_i}{Y_i} \right), \quad g = 1, 2, \dots, K$$

0에 가까울수록 귀무가설을 뒷받침

로그-순위 검정통계량

$$\chi^2 = (Z_1(\tau), \dots, Z_{K-1}(\tau)) \Sigma^{-1} (Z_1(\tau), \dots, Z_{K-1}(\tau))' \stackrel{H_0}{\sim} \chi_{K-1}^2$$

총화 로그-순위 검정

층화 로그-순위 검정

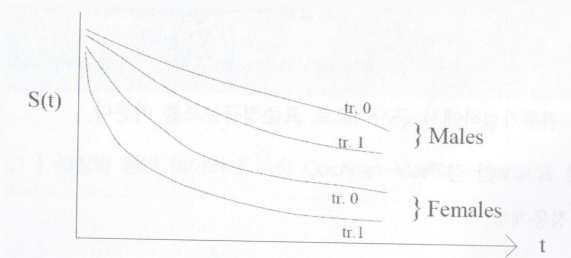
Example

두 개 그룹이 있고 각 그룹은 성별로 층화되어 있을 때 그룹별 생존함수 비교에 관심이 있는 경우

⇒ 성별 층을 보정한 생존함수를 추정하여 비교

층화 로그-순위 검정 (stratified log-rank test)

- K 개 그룹이 있고 각 집단은 S 개의 층으로 층화되어 있는 경우
- 각 층별로 로그-순위 검정통계량을 구한 후 통합하여 검정통계량을 구함



[그림 6.2] 성별(male/female)로 층화된 두 그룹(tr0/tr1) 생존함수의 관계

관심 가설

$$H_0 : h_{1s}(t) = h_{2s}(t) = \cdots = h_{Ks}(t), \quad s = 1, \dots, S, t \leq \tau$$

H_1 : 적어도 $t \leq \tau$ 의 일부분에서 한 개의 $h_{gs}(t)$ 는 다르다.

데이터 구조

$K(\geq 2)$ 개 모집단에서 발생한 데이터는 각 집단에서 S 개의 총을 갖는다.

로그-순위 검정통계량

$$Z_{j\cdot}(\tau) = \sum_{s=1}^S Z_{js}(\tau), \quad \hat{\sigma}_{jg\cdot} = \sum_{s=1}^S \hat{\sigma}_{jgs}$$

검정통계량

$$\chi^2 = (Z_{1\cdot}(\tau), \dots, Z_{K-1\cdot}(\tau)) \Sigma^{-1} (Z_{1\cdot}(\tau), \dots, Z_{K-1\cdot}(\tau))' \stackrel{H_0}{\sim} \chi_{K-1}^2$$

예제 6.6 : 총화의 예

- (a) 폐암 (lung cancer) 데이터를 여러 기관에서 수집한 경우 각 기관을 층으로 하여 총화에 따라 생존함수가 같다고 할 수 있는지 검정
survival 패키지에 내장된 데이터 `data(lung)`을 이용하여 총화 로그-순위 검정 [예 6.12] 참조
- (b) 폐암 데이터에 대해 성별을 층으로 하여 총화에 따라 생존함수가 같다고 할수 있는지 검정 [프로그램 6.7] 참조