

통계패키지 이론 및 실습

일반화 선형모형

-
- 범주형 반응변수 뿐만 아니라 연속형 반응변수에 대한 회귀, ANOVA모형들을 포함하는 매우 광범위한 모형
 - $g(\mu) = \alpha + \beta_1 x_1 + \cdots + \beta_p x_p$, where $\mu = E(Y)$, $g(\cdot)$: link function
 - $\pi(x) = \alpha + \beta x$: 선형확률모형
 - 최소제곱추정량보다 더 작은 표준오차를 갖는 최대우도추정량이 존재할 수도 있음

예) 코골이와 심장질환

코골이	심장질환	
	예	아니오
전혀 안함	24	1355
가끔씩	35	603
거의 매일 저녁	21	192
매일 저녁	30	224

-
- $\text{logit}(\pi(x)) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) : \text{로지트 연결}$
 - $\log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha + \beta x$
 - $\text{probit}(\pi(x)) = \Phi^{-1}(\pi(x)) : \text{프로빗 연결}$
 - $\text{probit}(\pi(x)) = \alpha + \beta x$

DATA glm;

INPUT snoring disease total;

CARDS;

0 24 1379

2 35 638

4 21 213

5 30 254

;

RUN;

PROC GENMOD DATA=glm;

MODEL disease/total=snoring / dist=bin link=identity;

RUN;

PROC GENMOD DATA=glm;

MODEL disease/total=snoring / dist=bin link=logit;

RUN;

PROC GENMOD DATA=glm;

MODEL disease/total=snoring / dist=bin link=probit;

RUN;

- 포아송 회귀모형

- $Y \sim Poi(\mu)$

- $\log(\mu) = \alpha + \beta x$

- $\mu = e^{\alpha} (e^{\beta})^x$

```
DATA crab;  
INFILE 'crab.txt';  
INPUT color spine width satell weight;  
RUN;  
PROC GENMOD DATA=crab;  
      MODEL satell=width / dist=poi link=log;  
RUN;
```