



일반화선형모형 (2)

정보통계학과 김성수교수

✓ 학습목차

1

로지스틱 회귀모형 해석

2

프로빗 모형

3

로그선형모형-개수형자료

4

로그선형모형-율(rate) 분석

1

로지스틱 회귀모형 해석

로지스틱 회귀모형 : 이항자료

〈날다람쥐 Sugar Glider의 출현자료〉

p_no	occurr	con_metric	p_size_km
1	1	0.650	130.9
2	0	0.610	104.1
3	0	0.744	132.3
4	1	0.213	225.6
5	1	0.723	83.0
6	0	0.678	48.8
7	0	0.733	61.0
8	1	0.522	39.6
9	1	0.552	193.1
10	0	0.245	155.6

```
> glider <-  
  read.csv('c:/data/reg/sugar_glider_binomial.csv')  
> head(glider, 3)  
  p_no occur con_metric p_size_km  
1    1    1    0.650    130.9  
2    2    0    0.610    104.1  
3    3    0    0.744    132.3  
> logit_m2 <- glm(occurr~p_size_km,  
                  family=binomial(link=logit),  
                  data=glider)
```

반응변수 $y = \text{occur}$, 1=yes, 0=no 이므로 이항분포를 가정

로지스틱 회귀모형 : $\eta = \log \text{it}(\pi) = \log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1$

$$\pi = E(Y | x) = \Pr(y = 1 | x)$$

로지스틱 회귀모형 : 이항자료

```
> summary(logit_m2)
```

```
...
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.5541	-0.8980	-0.5157	0.8075	2.0394

Coefficients:


	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-2.528298	0.820251	-3.082	0.00205	**
p_size_km	0.021727	0.006893	3.152	0.00162	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 68.994 on 49 degrees of freedom
Residual deviance: 55.716 on 48 degrees of freedom
AIC: 59.716

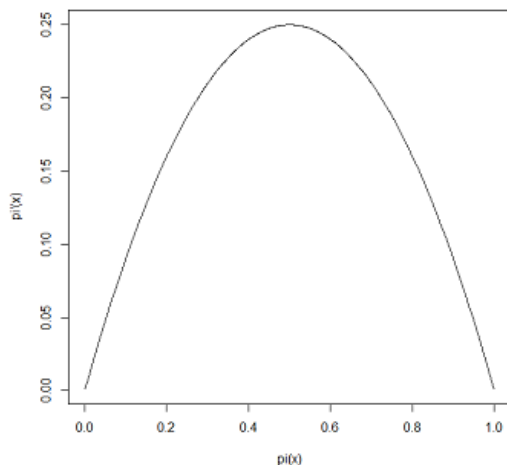
Number of Fisher Scoring iterations: 3


$$\log\left(\frac{\hat{\pi}(x)}{1 - \hat{\pi}(x)}\right) = -2.528 + 0.022 \times x$$

로지스틱 회귀의 선형 근사적 해석

x 의 변화에 따른 $\pi(x)$ 의 순간 변화량의 크기, 즉 x 값에서의 작은 변화가 성공률 $\pi(x)$ 에 미치는 영향의 크기

$$\pi'(x) = \frac{d\pi(x)}{dx} = \beta_1 \pi(x)[1 - \pi(x)]$$



- $\pi(x)$ 를 0.5가 되게 하는 x 값 근처에서 x 값이 변화할 때 $\pi(x)$ 의 순간 변화량은 약 $\beta_1/4$ 이 되어 x 의 변화가 주는 영향이 가장 큼.
- $\pi(x)$ 가 0 또는 1에 가까운 값이 되는 x 값 근처에서의 $\pi(x)$ 의 순간 변화량은 거의 0이 되어 x 의 변화에 따른 영향이 거의 없음.

〈로지스틱회귀모형 함수에서 x 의 변화에 따른 $\pi(x)$ 의 순간 변화량〉

로지스틱 회귀의 선형 근사적 해석

로지스틱 회귀모형: $\log\left(\frac{\hat{\pi}(x)}{1-\hat{\pi}(x)}\right) = -2.528 + 0.022 \times x$

- $\hat{\pi}(x) = 0.5$ 를 만족하는 구획의 크기

$$x_1 = -\frac{\hat{\beta}_0}{\hat{\beta}_1} = \frac{2.528}{0.022} = 114.9(km)$$

즉, 구획의 크기가 114.9km에서 115.9km 로
증가하면 Sugar glider가 출현할 확률

$\pi(x)$ 는 평균적으로 $\frac{\hat{\beta}_1}{4} = \frac{0.022}{4} = 0.006$ (0.6%p) 증가

하는 것으로 추정.

참고: $\log\left(\frac{\hat{\pi}(x)}{1-\hat{\pi}(x)}\right) = \hat{\beta}_0 + \hat{\beta}_1 x$

$$\hat{\pi}(x) = 0.5$$

$$\log\left(\frac{0.5}{1-0.5}\right) = \hat{\beta}_0 + \hat{\beta}_1 x = 0$$

$$x = -\frac{\hat{\beta}_0}{\hat{\beta}_1}$$

로지스틱 회귀와 승산비 및 성공확률

설명변수 X 가 x 일 때의 성공률 : $\pi(x) = P(Y=1|X=x)$

$X=x$ 일 때 성공일 승산(odds) : $\frac{\pi(x)}{1-\pi(x)}$

로지스틱회귀모형 $\log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x$ --- ①

: 로그-승산(log-odds)을 설명변수의 선형함수로 모형화 한 것

x 가 $x+1$ 로 증가하는 경우 :

$$\log\left(\frac{\pi(x+1)}{1-\pi(x+1)}\right) = \beta_0 + \beta_1(x+1) \quad \text{--- ②}$$

로지스틱 회귀와 승산비 및 성공확률

②-① 에서

$$\log\left(\frac{\pi(x+1) \times (1 - \pi(x))}{\pi(x) \times (1 - \pi(x+1))}\right) = \beta_1$$

양변에 로그의 역함수를 취하면

$$\frac{\pi(x+1) \times (1 - \pi(x))}{\pi(x) \times (1 - \pi(x+1))} = \frac{\pi(x+1)/(1 - \pi(x+1))}{\pi(x)/(1 - \pi(x))} = e^{\beta_1}$$

즉, e^{β_1} 은 x 가 $x+1$ 로 증가할 때 성공일 승산이 증가하는 비를 의미.

R 활용 : 승산비

```
> glider <- read.csv('c:/data/reg/sugar_glider_binomial.csv')
> logit_m2 <- glm(occure ~ p_size_km, family=binomial(link=logit), data=glider)
> exp(coef(logit_m2))
(Intercept)    p_size_km
  0.07979473  1.02196464
> exp(confint(logit_m2, parm="p_size_km", level=0.95))
Waiting for profiling to be done...
      2.5 %    97.5 %
1.009424 1.037535
```

결과해석 : 구획의 크기가 1km 증가할 때 Sugar Glider가 출현할 승산은 약 1.022배 증가하는 것으로 추정되며, 95% 신뢰수준에서 승산은 1.009~1.038배 사이에서 증가할 것으로 추정됨.

R 활용 : $\pi(x)$ 의 추정

$$\log\left(\frac{\hat{\pi}(x)}{1-\hat{\pi}(x)}\right) = -2.528 + 0.022 \times x$$

→
$$\hat{\pi}(x) = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 x)} = \frac{\exp(-2.528 + 0.022x)}{1 + \exp(-2.528 + 0.022x)}$$

x=150 에서 $\pi(x)$ 추정값 구하기

```
> x <- 150  
> predict(logit_m2, list(p_size_km=x), type="response")  
1  
0.6749669
```

※ $\pi(x)$ 의 신뢰구간 구하기 : 교재 참조

승산, 승산비 그리고 상대위험도

2×2 분할표의 확률분포

		Y	
		1	0
X	1	π_1	$1 - \pi_1$
	2	π_2	$1 - \pi_2$

$$OR = \frac{odds_1}{odds_2} = \frac{\pi_1 / (1 - \pi_1)}{\pi_2 / (1 - \pi_2)} = \frac{\pi_1 (1 - \pi_2)}{\pi_2 (1 - \pi_1)}$$

OR 값과 π_1 과 π_2 관계

① $OR = 1 \Leftrightarrow \pi_1 = \pi_2$

② $OR > 1 \Leftrightarrow \pi_1 > \pi_2$

③ $OR < 1 \Leftrightarrow \pi_1 < \pi_2$

승산, 승산비 그리고 상대위험도

2×2 분할표의 확률분포 (예)

		Y	
		1(폐암)	0(정상)
X	1(흡연군)	π_1	$1 - \pi_1$
	2(비흡연군)	π_2	$1 - \pi_2$

$$OR = \frac{odds_1}{odds_2} = \frac{\pi_1 / (1 - \pi_1)}{\pi_2 / (1 - \pi_2)} = \frac{\pi_1 (1 - \pi_2)}{\pi_2 (1 - \pi_1)}$$

상대위험도(RR; relative risk) : 성공률의 비 (위 예에서 관심질병의 발생률의 비)

$$RR = \frac{\pi_1}{\pi_2}$$

→ 예) $RR=2$ 라는 의미 : $X=1$ 에서의 관심질병의 발생률($=\pi_1$)이

$X=2$ 에서의 발생률($=\pi_2$) 보다 2배 높다 라는 의미

$OR=2$ 라는 의미 : $X=1$ 에서 관심질병이 발생할 승산이

$X=2$ 에서의 승산보다 2배 높다 라는 의미.

따라서 RR 보다는 해석이 명료하지는 않음.

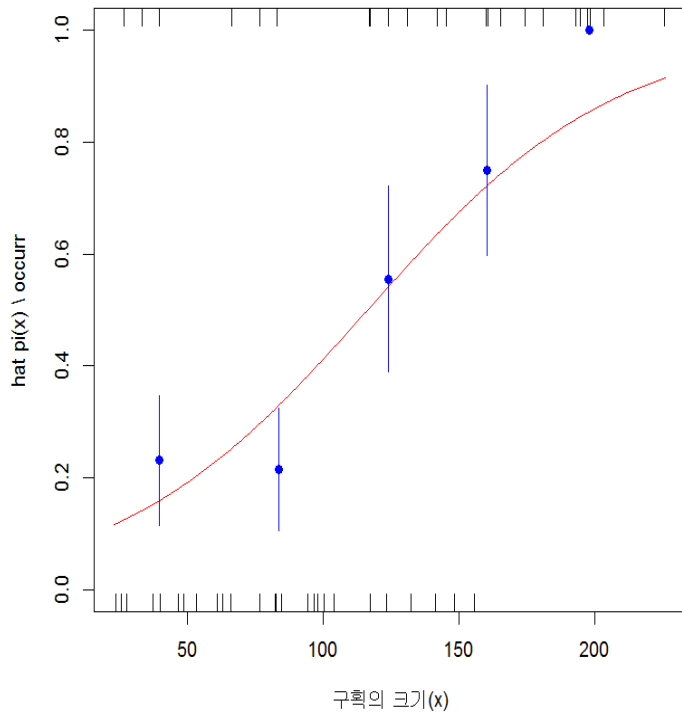
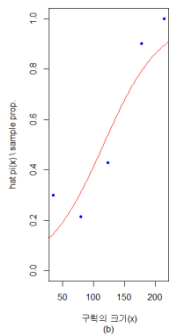
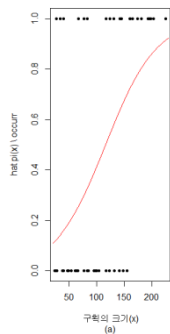
여기서 $\pi_1 \approx 0, \pi_2 \approx 0$ 인 경우

$$OR \approx RR$$

즉, 승산비를 상대위험도 처럼 해석이 가능

※(예제 7.2) 사례-대조 연구의 예 를 살펴보기 바람.

로지스틱 회귀모형 적합도의 시각적 효과



구획의 크기가 104~185km 사이에서는 모형의 적합도가 매우 높음을 알 수 있고, 205km 근처에서는 표본비율이 1이며 적합도가 높지 않음을 알 수 있다. 반면에 구획의 크기가 23.5~64.2 km 에서는 적합도가 어느 정도 있음을 알 수 있음.

R 코드 : 교재 참조

2 프로빗 모형

프로빗 모형

이항자료분석 : 로지스틱 회귀모형, 프로빗 모형

프로빗(Probit) 모형

① 반응변수의 확률분포: $Y \sim$ 이항분포 또는 베르누이분포

② 선형예측자: $\eta = x'\beta = \beta_0 + \beta_1x_1 + \dots + \beta_px_p$

③ 연결함수: $g(\pi) = \Phi^{-1}(\pi) = \beta_0 + \beta_1x_1 + \dots + \beta_px_p$

, π : 이항분포의 성공확률

$\Phi(\cdot)$: 표준정규분포의 누적분포함수

R 활용 : 프로빗 모형

```
> glider <- read.csv('c:/data/reg/sugar_glider_binomial.csv')  
> attach(glider)  
> probit_m <- glm(occurr~p_size_km, family=binomial(link=probit))
```

R 활용 : 프로빗 모형

```
> summary(probit_m)
```

```
...
```

```
Deviance Residuals:
```

Min	1Q	Median	3Q	Max
-1.5586	-0.9211	-0.5186	0.8041	2.0341

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.492825	0.460250	-3.244	0.001181	**
p_size_km	0.013023	0.003866	3.368	0.000757	***

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
(Dispersion parameter for binomial family taken to be 1)
```

```
Null deviance: 68.994  on 49  degrees of freedom
```

```
Residual deviance: 55.797  on 48  degrees of freedom
```

```
AIC: 59.797
```

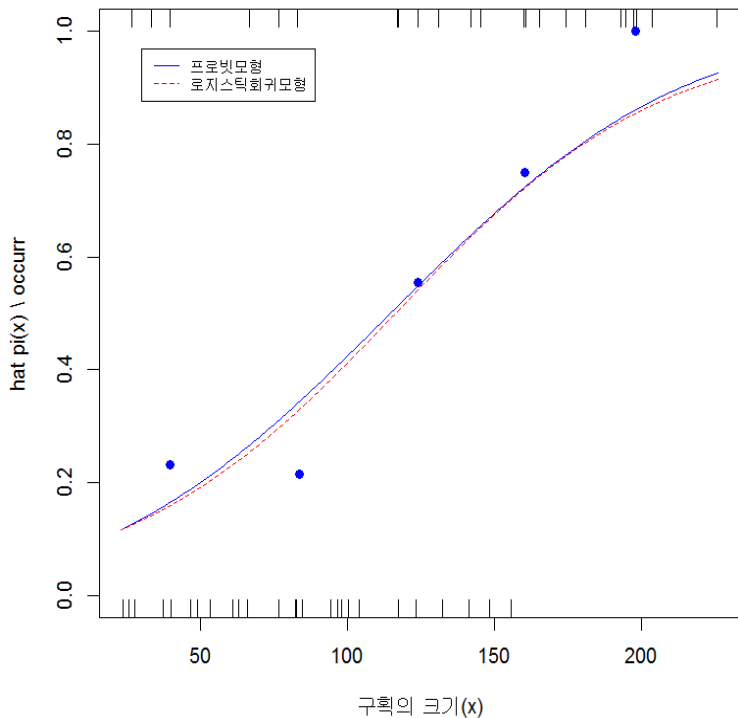
```
Number of Fisher Scoring iterations: 5
```

$$\Phi^{-1}(\hat{\pi}(x)) = \hat{\beta}_0 + \hat{\beta}_1 x = -1.493 + 0.013x$$

$$\hat{\pi}(x) = \Phi(\hat{\beta}_0 + \hat{\beta}_1 x) = \Phi(-1.493 + 0.013x)$$

모형이 적합

프로빗 모형과 로지스틱회귀모형 적합 결과



프로빗모형에서 $X=150\text{km}$ 일때, Sugar Glider가 출현할 확률 : 67.7%

```
> predict(probit_m, list(p_size_km=150),  
          type="response")
```

1
0.6774694

3 로그선형모형-개수형자료

로그선형모형

로그선형모형 : 개수형 자료(count data)를 분석할 때 자주 이용되는 모형

- ① 반응변수의 확률분포: $Y \sim$ 포아송분포
- ② 선형예측자: $\eta = x'\beta = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$
- ③ 연결함수: $g(\mu) = \log(\mu) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$

$\mu = E(Y)$: 포아송분포의 평균으로 사건의 평균발생건수를 나타냄.

- 로그연결함수를 평균에 대해 정리하면

$$\mu = \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)$$

가 되어 평균에 대한 승법모형(multiplicative model)이 됨.

- 그러므로 e^{β_i} 는 x_i 를 제외한 나머지 변수 $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p$ 들이 고정되었을 때, x_i 가 1단위 증가함으로써 발생하는 평균이 변화하는 비율(ratio)을 의미

R 활용 : 로그선형모형

< 고속도로 속도제한여부와 교통사고 건수 >

year	day	limit	y	year	day	limit	y
1961	1	no	9	1962	1	no	9
1961	2	no	11	1962	2	no	20
1961	3	no	9	1962	3	no	15
1961	4	no	20	1962	4	no	14
1961	5	no	31	1962	5	no	30
1961	6	no	26	1962	6	no	23
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

```
> library(MASS)
> data(Traffic)
> head(Traffic, 3)
```

	year	day	limit	y
1	1961	1	no	9
2	1961	2	no	11
3	1961	3	no	9

주요관심 내용 : 고속도로의 속도제한이 평균 사고건수에 어떤 영향을 주는가

분석모형 : $\log(\mu) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_{92} x_{92} + \beta_{93} x_{93}$

$$x_1 = \begin{cases} 0 & \text{no} \\ 1 & \text{yes} \end{cases}, x_i = \begin{cases} 1 & \text{day} = i \\ 0 & \text{나머지} \end{cases}, i = 2, 3, \dots, 92, x_{93} = \begin{cases} 0 & \text{year} = 1961 \\ 1 & \text{year} = 1962 \end{cases}$$

R 활용 : 로그선형모형

```
> Traffic$day <- as.factor(Traffic$day)
> Traffic$year <- as.factor(Traffic$year)
> log_m <- glm(y~limit+day+year, family=poisson(link=log), data=Traffic)
> summary(log_m)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	2.20984	0.23632	9.351	< 2e-16	***
limityes	-0.28424	0.04300	-6.610	3.86e-11	***
day2	0.54362	0.29633	1.834	0.066584	.
day3	0.28768	0.31180	0.923	0.356197	
day91	0.37539	0.31528	1.191	0.233800	
day92	0.64109	0.29876	2.146	0.031888	*
year1962	-0.02539	0.03458	-0.734	0.462927	

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 625.25 on 183 degrees of freedom

Residual deviance: 107.11 on 90 degrees of freedom

AIC: 1185.1

Number of Fisher Scoring iterations: 4

유의하지 않음

107.11/90=1.19 로 모형 적합

R 활용 : 로그선형모형

```
> log_m1 <- glm(y~limit+day, family=poisson(link=log), data=Traffic)
> summary(log_m1)
```

Coefficients: $\log(\hat{\mu}) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_{92} x_{92} = 2.20 - 0.30x_1 + 0.54x_2 + \dots + 0.65x_{92}$

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	2.19722	0.23570	9.322	< 2e-16 ***
limityes	-0.29627	0.03978	-7.448	9.46e-14 ***
day2	0.54362	0.29633	1.834	0.066584 .
...				
day91	0.38232	0.31515	1.213	0.225077
day92	0.64803	0.29862	2.170	0.030004 *

(Dispersion parameter for poisson family taken to be 1)

최종모형

Null deviance: 625.25 on 183 degrees of freedom
Residual deviance: 107.64 on 91 degrees of freedom
AIC: 1183.6
Number of Fisher Scoring iterations: 4

107.64/91=1.18 로 더 작아짐

AIC 값도 더 작아짐

로그선형모형의 해석

속도제한여부(limit)의 회귀계수 추정치 $\hat{\beta}_1$ 에 대한 해석

특정일 ($x_j, j=2, \dots, 92$)에

속도제한을 하지 않은 경우($x_1=0$)

평균사고건수($=\mu_0$)의 로그추정치 :

$$\log(\hat{\mu}_0) = \hat{\beta}_0 + 0 + \hat{\beta}_j x_j = 2.20 + 0 + \hat{\beta}_j x_j$$

속도제한을 한 경우($x_1=1$) 평균사고건수($=\mu_1$)의 로그추정치 :

$$\log(\hat{\mu}_1) = \hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_j x_j = 2.20 - 0.30 + \hat{\beta}_j x_j$$

$$\Rightarrow \log\left(\frac{\hat{\mu}_1}{\hat{\mu}_0}\right) = \hat{\beta}_1 = -0.30$$

$$\Rightarrow \frac{\hat{\mu}_1}{\hat{\mu}_0} = e^{-0.30} = 0.74$$

\Rightarrow 즉, 고속도로에서 속도제한을 했을 때의 평균사고건수는
속도제한을 하지 않았을 때의 평균사고건수의 74% 수준으로
감소

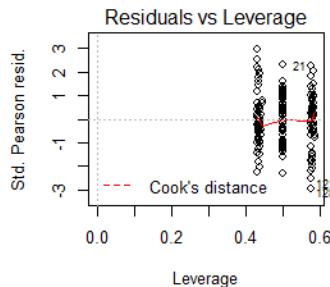
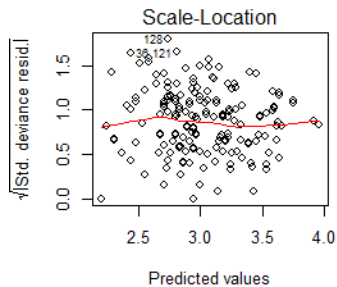
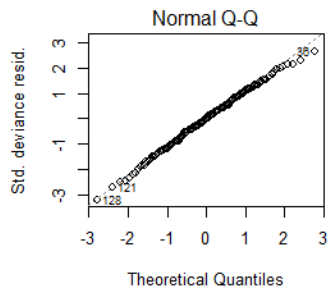
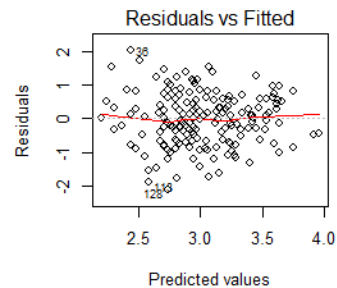
$\frac{\mu_1}{\mu_0} = e^{\beta_1}$ 의 추정치와 95% 신뢰구간

```
> exp(coef(log_m1, parm="limit"))  
(Intercept)  limityes      day2  
9.0000000    0.7435897    1.7222222  
...  
> exp(confint(log_m1, parm="limityes",  
  level=0.95))  
Waiting for profiling to be done...  
      2.5 %    97.5 %  
0.6877111 0.8037687
```

고속도로에서 속도제한을 실시하면 평균사고건수
가 26% 정도 감소하며 95% 신뢰수준에서 많게는
31% 적게는 20% 정도 감소하는 것으로 추정됨

로그선형모형의 잔차진단

```
> par(mfrow=c(2,2))  
> plot(log_m1)
```



네 가지 잔차 그림 모두 잔차 분포가 정규성을 크게 위배하지 않음을 보여주고 있음.

4 로그선형모형-율(rate) 분석

로그선형모형 - 율(rate) 분석

- 관심사건의 발생률이 매우 작은 경우, 전체 관측건수가 매우 많아 관심사건을 일정 수 이상 관측할 수 있음. 이 경우, 관심사건의 수 Y 는 포아송분포에 근사한다고 가정.
- λ 를 관심사건의 발생율, N 을 전체 관측건수, Y 를 관심사건의 수라고 할 때, λ 의 분석 모형

① 반응변수의 확률분포: $Y \sim \text{poisson}(N\lambda)$

② 선형예측자: $\eta = x'\beta = \beta_0 + \beta_1x_1 + \dots + \beta_px_p$

③ 연결함수: $g(\mu) = \log(N\lambda) = \beta_0 + \beta_1x_1 + \dots + \beta_px_p$, 즉

$$\log(\lambda) = -\log(N) + \beta_0 + \beta_1x_1 + \dots + \beta_px_p$$

, $\log(N)$ 은 모형에 주어지는 상수이며 이와 같은 항을 상쇄항(오프셋:offset)이라 함.

R 활용 : 율(rate) 분석

나이와 지역별 흑색종 발병 자료

age	region	cases	total	age	region	cases	total
-35	south	64	1074246	-35	north	61	2880262
35-44	south	75	220407	35-44	north	76	564535
45-54	south	68	198119	45-54	north	98	592983
55-64	south	63	134084	55-64	north	104	450740
65-74	south	45	70708	65-74	north	63	270908
75+	south	27	34233	75+	north	80	161850

관심 내용 : 흑색종 발병률이
나이(age)와 거주지역(region)에
따라 차이가 있는가

분석모형 λ_x : 흑색종 발병률, N_x : 전체 조사건수

$$\log(\lambda_x) = -\log(N_x) + \beta_0 + \beta_1 x_1 + \dots + \beta_5 x_5 + \beta_6 x_6$$

$$, x_1 = \begin{cases} 1 & \text{age} = 35-44 \\ 0 & \text{나머지} \end{cases}, x_2 = \begin{cases} 1 & \text{age} = 45-54 \\ 0 & \text{나머지} \end{cases},$$

$$x_3 = \begin{cases} 1 & \text{age} = 55-64 \\ 0 & \text{나머지} \end{cases}, x_4 = \begin{cases} 1 & \text{age} = 65-74 \\ 0 & \text{나머지} \end{cases},$$

$$x_5 = \begin{cases} 1 & \text{age} = 75+ \\ 0 & \text{나머지} \end{cases}, x_6 = \begin{cases} 1 & \text{region} = \text{south} \\ 0 & \text{region} = \text{north} \end{cases}$$

R 활용 : 로그선형모형

```
> melanoma <- read.csv('c:/data/reg/melanoma.csv')
```

```
> melanoma
```

	age	region	cases	total
1	35-44	south	75	220407
2	45-54	south	68	198119
3	55-64	south	63	134084
4	65-74	south	45	70708
5	75+	south	27	34233
6	<35	south	64	1074246
7	35-44	north	76	564535
8	45-54	north	98	592983
9	55-64	north	104	450740
10	65-74	north	63	270908
11	75+	north	80	161850
12	35	north	61	2880262

```
> log.rt <- glm(cases~age+region-offset(log(total)),  
                family=poisson(link=log), data=melanoma)
```

R 활용 : 로그선형모형

```
> summary(log.rt)
```

```
...
```

```
Deviance Residuals:
```

1	2	3	4	5	6	7	8
0.4780	-0.4273	-0.5302	0.7469	-1.3610	0.8686	-0.4581	0.3667
9	10	11	12				
0.4279	-0.5932	0.8904	-0.8283				

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-10.65831	0.09518	-111.97	<2e-16 ***
age35-44	1.79737	0.12093	14.86	<2e-16 ***
age45-54	1.91309	0.11844	16.15	<2e-16 ***
age55-64	2.24180	0.11834	18.94	<2e-16 ***
age65-74	2.36572	0.13152	17.99	<2e-16 ***
age75+	2.94468	0.13205	22.30	<2e-16 ***
regionsouth	0.81948	0.07103	11.54	<2e-16 ***

```
---
```

```
(Dispersion parameter for poisson family taken to be 1)
```

```
Null deviance: 895.8197 on 11 degrees of freedom
```

```
Residual deviance: 6.2149 on 5 degrees of freedom
```

```
AIC: 92.44
```

```
Number of Fisher Scoring iterations: 4
```

$$\begin{aligned}\log(\hat{\lambda}_x) &= -\log(N_x) + \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_5 x_5 + \hat{\beta}_6 x_6 \\ &= -\log(N_x) - 10.66 + 1.80x_1 + \dots + 2.94x_5 + 0.82x_6\end{aligned}$$

로그선형모형의 해석

거주지역 효과인 β_6 추론 : e^{β_6} 의 추정치와 95% 신뢰구간

```
> exp(coef(log_rt))
(Intercept)  age35-44  age45-54  age55-64  age65-74  age75+
2.350472e-05 6.033788e+00 6.773973e+00 9.410278e+00 1.065175e+01 1.900457e+01
regionsouth
2.269330e+00
> exp(confint(log_rt, parm="regionsouth",level=0.95))
Waiting for profiling to be done...
      2.5 %   97.5 %
1.973244 2.607037
```

남쪽지역의 흑색종 발병률은 북쪽지역 보다 $e^{\hat{\beta}_6} = 2.27$ 배 높은 것으로 추정되며,
95% 신뢰수준에서 작게는 1.97배 크게는 2.61배 높은 것으로 추정됨.



다음시간 안내

14강. SAS와 SPSS 활용 (일반화선형모형)