15.10.2015 / ex 19, ex 5/5, P2E 2004 (1)

a)

	Group 1 ($x_i = 1$)	Group 0 ($x_i = 0$)	Total
Sick	A	B	$n_{0\cdot} = A + B$
Healthy	C	D	$n_{1\cdot} = C + D$
Total	$n_{\cdot 1} = A + C$	$n_{\cdot 0} = B + D$	$n = A + B + C + D$

Table 1. Number of people with and without diabetes by gender

Gender	Male	Female	Total
Diseased	377	336	713
Healthy	17864	20099	37963
Total	18241	20435	38676

under 25

BMI	< 25	≥ 25	Total
Diseased	90	623	713
Healthy	21689	16274	37963
Total	21779	16897	38676

$Y_i \in \{0, 1\}$ $\quad \eta_i = \alpha + \beta X_i = \alpha + \beta \mathbb{1}\{X_i = 1\}$

$X_i \in \{0, 1\}$ $\quad \Pi_i = g^{-1}(\eta_i) = \begin{cases} \Pi(0), & \text{if } X_i = 0 \\ \Pi(1), & \text{if } X_i = 1 \end{cases}$

$\Pi(0) = P(Y = 0 \mid X = 0)$ $\quad A \sim Binom(n = A + C, \Pi(1))$

$\Pi(1) = P(Y = 0 \mid X = 1)$ $\quad B \sim Binom(n = B + D, \Pi(0))$

$$l(\Pi(0)) = \log\binom{B+D}{B} + B \log \Pi(0) + D \log(1 - \Pi(0))$$

$$\frac{\partial l(\Pi(0))}{\partial \Pi(0)} = \frac{B}{\Pi(0)} - \frac{D}{1 - \Pi(0)} \iff l'(\Pi(0)) = 0$$

$$\iff \Pi(0) = \frac{B}{B + D} \text{ and same way } \Pi(1) = \frac{A}{A + C} \Rightarrow$$

$$\Rightarrow J = -\frac{\partial^2 l(\Pi(0))}{\partial \Pi(0) \partial \Pi(0)} = \frac{B}{\Pi^2(0)} + \frac{D}{(1 - \Pi(0))^2} \approx (Var(\hat{\Pi}(0)))^{-1}$$

$$OR = \frac{\frac{\Pi(1)}{1 - \Pi(1)}}{\frac{\Pi(0)}{1 - \Pi(0)}} = \frac{\Pi(1)(1 - \Pi(0))}{\Pi(0)(1 - \Pi(1))}$$

$$\hat{OR} = \frac{\frac{A}{A+C} \cdot \frac{D}{B+D}}{\frac{B}{B+D} \cdot \frac{C}{A+C}} = \frac{A \cdot D}{B C}$$

$$\theta(j) = \log\left(\frac{\Pi(j)}{1 - \Pi(j)}\right), \quad j \in \{0, 1\}$$

$$\hat{\theta}(j) = \log\left(\frac{\hat{\Pi}(j)}{1 - \hat{\Pi}(j)}\right) = [j = 1] = \log\left(\frac{\frac{A}{A+C}}{\frac{C}{A+C}}\right) =$$

$$= \log A - \log C$$

$$Var(\hat{\theta}(j)) = Var(\Pi(j)) \left(\frac{\partial h(\Pi(j))}{\partial \Pi(j)}\right)^2$$

b) $\theta(j) = \log\left(\frac{\Pi(j)}{1 - \Pi(j)}\right) = h(\Pi(j))$ with $j = 1$

$$\frac{\partial h(\Pi(1))}{\partial \Pi(1)} = \frac{1}{\frac{\Pi(1)}{1 - \Pi(1)}} \left(\frac{1}{1 - \Pi(1)} + \frac{\Pi(1)}{(1 - \Pi(1))^2}\right) =$$

$$= \frac{1}{\Pi(1)(1 - \Pi(1))}$$

we have found $Var(\hat{\Pi}(1)) \approx \left(\frac{A}{\Pi^2(1)} + \frac{C}{(1 - \Pi(1))^2}\right)^{-1}$

w.r.t Delta method: $Var(\hat{\theta}(1)) = \left(\frac{A}{\Pi^2(1)} + \frac{C}{(1 - \Pi(1))^2}\right)^{-1} \times$

$$\times \left(\frac{1}{\Pi(1)(1 - \Pi(1))}\right)^2 \approx \left[\hat{\Pi}(1) = \frac{A}{A+C}\right] \approx$$

$$\approx \left(\frac{A(A+C)^2}{A^2} + \frac{C(A+C)^2}{C^2}\right)^{-1} \times \frac{1}{\left(\frac{A}{A+C}\right)^2 \left(\frac{C}{A+C}\right)^2} = (A+C)^{-2}\left(\frac{1}{A} + \frac{1}{C}\right)^{-1} \frac{(A+C)^4}{A^2 C^2} =$$

$$= (A+C)^2 \frac{A C}{(A+C) A^2 C^2} = \frac{1}{C} + \frac{1}{A}$$

in the same way we get:

$$Var(\theta(0)) \approx \frac{1}{B} + \frac{1}{D} \Rightarrow$$

c) $Var(\log \hat{OR}) = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D}$

$$\log \hat{OR} = \log\left(\frac{\hat\pi(1)}{1-\hat\pi(1)} \Big/ \frac{\hat\pi(0)}{1-\hat\pi(0)}\right) = \left[\begin{matrix}\hat\pi(1) = \frac{A}{A+C} \\ \hat\pi(0) = \frac{B}{B+D}\end{matrix}\right] =$$

$$= \left[\log\frac{\pi(1)}{1-\pi(1)} = \theta(1)\right] =$$

$$= \hat\theta(1) - \hat\theta(0)$$

$$var(\log\hat{OR}) = var(\hat\theta(1) - \hat\theta(2)) = Var(\hat\theta(1)) + Var(\hat\theta(2)) =$$

$$\approx \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D} = se^2$$

$$\log OR \in \log\hat{OR} \pm z_{\frac{\alpha}{2}}\sqrt{Var(\log\hat{OR})} =$$

$$= \underline{\log\hat{OR} \pm z_{\frac{\alpha}{2}} \cdot se}$$

$$OR \in \exp(\log\hat{OR} \pm z_{\frac{\alpha}{2}} \cdot se] =$$

$$= \hat{OR}\exp(\pm z_{\frac{\alpha}{2}} se) = \left[\begin{matrix}\alpha = 0.05 \Rightarrow \\ z_{\frac{\alpha}{2}} = 1.96\end{matrix}\right]$$

$$d) = \hat{OR}\exp(\pm 1.96 se) \quad , i \in I, se]$$

$$BT = \sum_{ij}\frac{n_{ij} - \hat n_{ij}}{\hat n_{ij}} = \left[\hat n_{ij} = \frac{n_{i.} n_{j.}}{n}\right] =$$

$$= \sum_{ij}\frac{n_{ij} - \frac{n_{i.} n_{j.}}{n}}{\frac{n_{i.} n_{j.}}{n}} \sim \chi^2_{(\|I\|-1)(\|J\|-1)} \quad) =$$

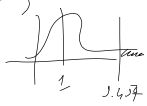
$$\overset{d}{=} \chi^2_1$$

$$BT = 9.497 >> E\{\chi^2_1\} = 1 \Rightarrow$$

extremely small p-value

p-value = 0.0021 ⇒

reject H_0!



$H_0: \pi(0) = \pi(1)$

$$\hat{OR} = \frac{AD}{CB} = \frac{377}{17864} \times \frac{20099}{336} = \underline{1.26}$$

$$se^2_{\log\hat{OR}} = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D} =$$

$$= 0.0057$$

$$OR \in 1.26\exp(\pm z_{\frac{\alpha}{2}}\sqrt{0.0057}) =$$

f)

$$\frac{\frac{\pi(1)}{1-\pi(1)}}{\frac{\pi(0)}{1-\pi(0)}} = \frac{\frac{1}{\frac{1}{\pi(1)} - 1}}{\frac{1}{\frac{1}{\pi(0)} - 1}} \overset{\pi(1),\pi(0) > 0}{=} 1 \Leftrightarrow$$

$$\Leftrightarrow \boxed{\pi(1) = \pi(0)}$$

$$logit(\pi(i)) = \log\left(\frac{\pi(i)}{1-\pi(i)}\right) = \log(odds(i))$$

$$\frac{\exp(logit(\pi(1)))}{\exp(logit(\pi(0)))} = OR = \left[\begin{matrix}logit(\pi(i)) = g(\pi(i)) = \eta_i \\ = \alpha + \beta x_i\end{matrix}\right] =$$

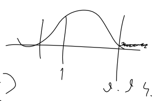
$$\frac{\exp(\alpha + \beta \cdot 1)}{\exp(\alpha + \beta \cdot 0)} = \exp(\beta) = 1 \Leftrightarrow$$

$$\Leftrightarrow \beta = 0$$

$$\begin{cases} H_0: \beta = 0 \\ H_1: \beta \neq 1.\end{cases} \qquad W = \frac{\hat\beta - 0}{\sqrt{Var(\hat\beta)}} = 9.445 \sim \chi^2_1$$

$$9.445 >> E\{\chi^2_1\} = 1 \Rightarrow low$$

p-value = 0.0021



⇒ reject H_0 ⇒ $\beta \neq 0$ ⇒

odds ratio is not 1

$$h). \ \hat{OR} = \frac{A}{C}\frac{D}{B} = \underline{0.1084}$$

$$Var_{\log\hat{OR}} = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D} = 0.0128$$

$$OR \in \hat{OR}\exp(\pm z_{\frac{\alpha}{2}} se_{\log OR}) \overset{\alpha=0.05}{=} 0.1084\exp(\pm 1.96\sqrt{0.0128}) =$$

$$= (0.0868, 0.1353). \quad \boxed{OR_{HB} << 1}$$

g) for g) see code on webpage

exam 2004, Problem (2)

		Relatives with diabetes			
		Yes		No	
		Age at diagnosis		Age at diagnosis	
		< 45	≥ 45	< 45	≥ 45
Dependent of injection of insulin	Yes	6	6	16	8
	No	1	36	2	48

$y_{jkl} \in \{0, \ldots N\}_{jkl}$ / with deceased relatives / without them

$j \in J = \{1, 2\}$ — 1: dep, 2: no dep on insulin. $k \in K = \{1, 2\}$

$l \in L = \{1, 2\}$ — younger, older

a) ① $y_{kl} \sim Binom(n_{kl}, \pi_{kl})$ belongs to Exp. family

$\eta_{kl} = logit(\pi_{kl})$, $\pi_{kl} = \frac{e^{\eta_{kl}}}{1 + e^{\eta_{kl}}} \in \{0, 1\}$

② $g(\pi_{kl}) = logit(\pi_{kl})$ – canonical link

③ $\eta_{kl} = \beta_0 + \beta_1 age_{kl} + \beta_2 rel_{kl} + \beta_3 age_{kl} \cdot rel_{kl}$ — linear predictor

$$\begin{pmatrix} \eta_{11} \\ \eta_{12} \\ \eta_{21} \\ \eta_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} =$$

corner point parametrization

$$= \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \beta_0 + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \beta_1 + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \beta_2 + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \beta_3$$

---

b). $odds_{kl} = \frac{\pi_{kl}}{1 - \pi_{kl}} = \exp(\eta_{kl})$

$$\hat{OR} = \frac{\hat{odd}_{k_1 l_1} | k = c, l_1 = 1}{\hat{odd}_{k_2 l_2} | k = c, l_2 = 2} = \frac{\exp(\eta_{k_1 l_1})}{\exp(\eta_{k_2 l_2})} =$$

$$= \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 \cdot 0)}{\exp(\hat{\beta}_0 + \hat{\beta}_1 \cdot 1)} = \exp(-\hat{\beta}_1) = \exp(3.78) \approx 43.82$$

$\hat{OR} = \exp(-\hat{\beta}_1)$ $\alpha = 0.05$

$\beta_1 \in \hat{\beta}_1 \pm z_{\frac{\alpha}{2}} \sqrt{var(\hat{\beta}_1)} =$

$= -3.78 \pm 1.96 \cdot 0.68 = (-5.12, -2.45)$

$OR = \exp(-\beta_1) \in \exp(-3.78 \pm 1.96 \times 0.68) =$

$= [11.52, 167.08]$

c) $\Delta = 2(\ell(\tilde{\pi}) - \ell(\hat{\pi})) = [$

Binomial data $] = 2 \sum \log \binom{n_{kl}}{y_{kl}} +$

$+ y_{kl} \log \tilde{\pi}_{kl} + (n_{kl} - y_{kl}) \log(1 - \tilde{\pi}_{kl}) -$

$- \log \binom{n_{kl}}{y_{kl}} - y_{kl} \log \hat{\pi}_{kl} - (n_{kl} - y_{kl}) \log(1 - \hat{\pi}_{kl}) =$

$= [$ perfect fit: $n_{kl} \tilde{\pi}_{kl} = y_{kl}$, $\tilde{\pi}_{kl} = \frac{y_{kl}}{n_{kl}}] = 2\sum$

$$= y_{kl} \log \left(\frac{y_{kl} / n_{kl}}{\hat{\pi}_{kl}} \right) + (n_{kl} - y_{kl}) \log \left(\frac{\frac{n_{kl} - y_{kl}}{n_{kl}}}{1 - \hat{\pi}_{kl}} \right) =$$

$k = 2, l = 1$:

$$= \sum_{kl} \Delta_{kl} \qquad \Delta_{21} = 2 \left(y_{21} \log \left(\frac{y_{21}}{n_{21} \hat{\pi}_{21}} \right) + \right.$$

$$\left. + (n_{21} - y_{21}) \log \left(\frac{n_{21} - y_{21}}{n_{21}(1 - \hat{\pi}_{21})} \right) \right) = \left[\begin{matrix} n_{21} = 18 \\ y_{21} = 16 \end{matrix} \right]$$