ניתוח נתונים בינאריים - פתרון לתרגיל מס׳ 1

1. א. מונה של z:

$$\frac{n_{01}}{n_{0\cdot}} - \frac{n_{11}}{n_{1\cdot}} = \frac{1}{n_{0\cdot}n_{1\cdot}}\left(n_{01}n_{1\cdot} - n_{11}n_{0\cdot}\right)$$

$$= \frac{1}{n_{0\cdot}n_{1\cdot}}\left[n_{01}(n_{10}+n_{11}) - n_{11}(n_{00}+n_{01})\right]$$

$$= \frac{1}{n_{0\cdot}n_{1\cdot}}\left[n_{01}n_{10} + n_{01}n_{11} - n_{11}n_{00} - n_{11}n_{01}\right]$$

$$= \frac{-1}{n_{0\cdot}n_{1\cdot}}\left[n_{00}n_{11} - n_{01}n_{10}\right]$$

מכנה של z (בלי בתוך השורש):

$$\frac{n_{\cdot 0}}{n_{\cdot\cdot}}\frac{n_{\cdot 1}}{n_{\cdot\cdot}}\left(\frac{1}{n_{0\cdot}} + \frac{1}{n_{1\cdot}}\right)$$

$$= \frac{n_{\cdot 0}n_{\cdot 1}}{n_{\cdot\cdot}^2}\left(\frac{n_{0\cdot}+n_{1\cdot}}{n_{0\cdot}n_{1\cdot}}\right) = \frac{1}{n_{\cdot\cdot}}\frac{n_{\cdot 0}n_{\cdot 1}}{n_{0\cdot}n_{1\cdot}}$$

($n_{0\cdot}+n_{1\cdot} = n_{\cdot\cdot}$)

לכן:

$$z^2 = \left(\frac{1}{n_{0\cdot}n_{1\cdot}}\right)^2 (n_{00}n_{11} - n_{01}n_{10})^2 \Big/ \left\{\left(\frac{1}{n_{\cdot\cdot}}\right)\left(\frac{n_{\cdot 0}n_{\cdot 1}}{n_{0\cdot}n_{1\cdot}}\right)\right\}$$

This proves that
(z')^2 = C

$$= n_{\cdot\cdot}(n_{00}n_{11} - n_{01}n_{10})^2 \big/ (n_{0\cdot}n_{1\cdot}n_{\cdot 0}n_{\cdot 1}) = \chi^2.$$

ב. ראו הדף הבא (שם הסימון הוא $n_{11}, n_{12}, n_{21}, n_{22}$ במקום הסימון $n_{00}, n_{01}, n_{10}, n_{11}$ שאני משתמש בו בשיעורים כעת).

Here I used the notation N_{ij} for what in denoted in class by E_{ij} .

תרגיל 1, שאלה 1 (ב)

$N_{11} + N_{12} = n_{1\cdot}$. ולפי $n_{11} + n_{12} = n_{1\cdot}$. לפיכך

$n_{12} - N_{12} = -(n_{11} - N_{11})$.

באופן דומה:

$n_{21} - N_{21} = -(n_{11} - N_{11})$

$n_{22} - N_{22} = -[-(n_{11} - N_{11})] = n_{11} - N_{11}$.

כלומר $(n_{ij} - N_{ij})^2$ הוא אותו דבר לכל i,j.

לכן ניתן לרשום

$$\chi^2 = \sum_{i,j=1}^{2} \frac{(n_{ij} - N_{ij})^2}{N_{ij}}$$

$$= (n_{11} - N_{11})^2 \left[\frac{1}{N_{11}} + \frac{1}{N_{12}} + \frac{1}{N_{21}} + \frac{1}{N_{22}} \right]$$

$$= (n_{11} - N_{11})^2 \left[\frac{N_{11} + N_{12}}{N_{11} N_{12}} + \frac{N_{21} + N_{22}}{N_{21} N_{22}} \right]$$

$$= (n_{11} - N_{11})^2 \left[\frac{n_{1\cdot}}{N_{11} N_{12}} + \frac{n_{2\cdot}}{N_{21} N_{22}} \right]$$

$$= n_{\cdot\cdot}^2 (n_{11} - N_{11})^2 \left[\frac{n_{1\cdot}}{n_{1\cdot} n_{\cdot 1} n_{1\cdot} n_{\cdot 2}} + \frac{n_{2\cdot}}{n_{2\cdot} n_{\cdot 1} n_{2\cdot} n_{\cdot 2}} \right]$$

$$= (n_{11} - N_{11})^2 \Big/ \left[\frac{n_{1\cdot} n_{2\cdot} n_{\cdot 1} n_{\cdot 2}}{n_{\cdot\cdot}^3} \right] .$$

This proves A=B

הוכיח ש- $B = C$

$$B = \frac{(n_{11} - E_{11})^2}{\frac{n_{0\cdot} n_{1\cdot} n_{\cdot 0} n_{\cdot 1}}{n_{\cdot\cdot}^3}} = \frac{\left(n_{11} - \frac{n_{1\cdot} n_{\cdot 1}}{n_{\cdot\cdot}}\right)^2}{\frac{n_{0\cdot} n_{1\cdot} n_{\cdot 0} n_{\cdot 1}}{n_{\cdot\cdot}^3}}$$

$$= \frac{(n_{11} n_{\cdot\cdot} - n_{1\cdot} n_{\cdot 1})^2 / n_{\cdot\cdot}^2}{n_{0\cdot} n_{1\cdot} n_{\cdot 0} n_{\cdot 1} / n_{\cdot\cdot}^3}$$

$$= \frac{n_{\cdot\cdot}(n_{11} n_{\cdot\cdot} - n_{1\cdot} n_{\cdot 1})^2}{n_{0\cdot} n_{1\cdot} n_{\cdot 0} n_{\cdot 1}}$$

נסמן

$$n_{11} n_{\cdot\cdot} - n_{1\cdot} n_{\cdot 1}$$

$$= n_{11}(n_{00} + n_{01} + n_{10} + n_{11}) - (n_{10} + n_{11})(n_{01} + n_{11})$$

$$= n_{11} n_{00} + n_{11} n_{01} + n_{10} n_{11} + n_{11}^2$$

$$- [n_{10} n_{01} + n_{11} n_{01} + n_{10} n_{11} + n_{11}^2]$$

$$= n_{00} n_{11} - n_{01} n_{10}$$

ולכן כנדרש כי $B = C$.

.2

לוח של odds ratio ω

$\pi_{B|\bar{A}}$	R = 1.0	1.5	2.0	2.5	3.0
0.01	1.00	1.51	2.02	2.54	3.06
0.05	1.00	1.54	2.11	2.71	3.35
0.10	1.00	1.59	2.25	3.00	3.86
0.15	1.00	1.65	2.43	3.40	4.64
0.20	1.00	1.71	2.67	4.00	6.00

רואים את הקירבה
בין ω ל- R כאשר
$\pi_{B|\bar{A}}$ נמוך והפער
בינהם כאשר $\pi_{B|\bar{A}}$
גבוה יותר

בס"ד

3.

$$\log \hat{R} = \log \hat{\pi}_{B|A} - \log \hat{\pi}_{B|\bar{A}}$$

עבור $g(p)=\log p$ יש לנו $g'(p) = 1/p$. לכן לפי שיטת דלתא

$$Var(\log \hat{\pi}_{B|A}) \doteq \left(\frac{1}{\pi_{B|A}}\right)^2 Var(\hat{\pi}_{B|A})$$

Here again, the notation differs from what I used in class, but it should be clear.

$$= \left(\frac{1}{\pi_{B|A}}\right)^2 \frac{\pi_{B|A}(1-\pi_{B|A})}{n_{1\cdot}}$$

$$= \frac{1-\pi_{B|A}}{\pi_{B|A}\, n_{1\cdot}} = \frac{1}{n_{1\cdot}} \frac{\pi_{\bar{B}|A}}{\pi_{B|A}}$$

$$\hat{Var}(\log \hat{\pi}_{B|A}) = \frac{1}{n_{1\cdot}} \frac{n_{10}/n_{1\cdot}}{n_{11}/n_{1\cdot}} = \frac{1}{n_{1\cdot}} \frac{n_{10}}{n_{11}}$$

באופן דומה

$$Var(\log \hat{\pi}_{B|\bar{A}}) = \frac{1}{n_{0\cdot}} \frac{\pi_{\bar{B}|\bar{A}}}{\pi_{B|\bar{A}}}$$

$$\hat{Var}(\log \hat{\pi}_{B|A}) = \frac{1}{n_{0\cdot}} \frac{n_{00}}{n_{01}}$$

לכן

$$Var(\log \hat{R}) = \frac{1}{n_{0\cdot}} \frac{\pi_{\bar{B}|\bar{A}}}{\pi_{B|\bar{A}}} + \frac{1}{n_{1\cdot}} \frac{\pi_{\bar{B}|A}}{\pi_{B|A}}$$

$$\hat{Var}(\log \hat{R}) = \frac{1}{n_{0\cdot}} \frac{n_{00}}{n_{01}} + \frac{1}{n_{1\cdot}} \frac{n_{10}}{n_{11}}$$

5. מ"ס של R:

$$\hat{R} = \frac{180/200}{150/200} = 1.2 \qquad (= 1/0.8333)$$

$$\log \hat{R} = \log 1.2 = 0.1823$$

לפי הנ"ו בשלב הקודם

$$\widehat{Var}(\log \hat{R}) = \frac{1}{200}\,\frac{50}{150} + \frac{1}{200}\,\frac{20}{180} = 0.002222$$

$$\widehat{Var}(\log \hat{R})^{1/2} = \sqrt{0.002222} = 0.04714$$

רו"ס (ברמת 95%) של $\log R$

$$0.1823 \pm (1.96)(0.04714) \Rightarrow [0.08993, 0.2747]$$

רו"ס של R:

$$[e^{0.08993}, e^{0.2747}]$$

$$= [1.0941, 1.3162] = [(0.9140)^{-1}, (0.7598)^{-1}]$$

ניתן גם להשתמש ב SAS.

בס"ד

ר"ס ל ω:

$$\hat{\omega} = \frac{(50)(180)}{(150)(20)} = 3$$

$$\log \hat{\omega} = \log 3 = 1.0986$$

לפי הפיתוחים בכיתה,

$$\widehat{Var}(\log \hat{\omega}) = \frac{1}{50} + \frac{1}{150} + \frac{1}{20} + \frac{1}{180} = 0.0822$$

$$[\widehat{Var}(\log \hat{\omega})]^{1/2} = \sqrt{0.0822} = 0.2867$$

ר"ס ל $\log \omega$:

$$1.0986 \pm (1.96)(0.2867) \Rightarrow [0.5366, 1.6606]$$

ר"ס ל- ω:

$$[e^{0.5366}, e^{1.6606}] = [1.7102, 5.2626]$$

תואם עם SAS.