# 4. Contingency tables

## (a)

2/2 points (graded)

Even though logistic regression is formulated with continuous input data in mind, one can also try to apply it to categorical inputs. For example, consider the following set-up: We observe $n$ samples $Y_i \in \{0,1\}$, $i = 1,\ldots,n$, and covariates $X_i \in \{0,1\}$, $i = 1,\ldots,n$. Moreover, assume that given $X_i$, the $Y_i$ are independent.

First, let us apply regular maximal likelihood estimation. To this end, write

$$\begin{aligned} f_{00} = &\quad \frac{1}{n}\#\{i : X_i = 0 \text{ and } Y_i = 0\} \\ f_{01} = &\quad \frac{1}{n}\#\{i : X_i = 0 \text{ and } Y_i = 1\} \\ f_{10} = &\quad \frac{1}{n}\#\{i : X_i = 1 \text{ and } Y_i = 0\} \\ f_{11} = &\quad \frac{1}{n}\#\{i : X_i = 1 \text{ and } Y_i = 1\} \end{aligned}$$

and assume that $f_{00}, f_{01}, f_{10}, f_{11} > 0$. We can parametrize this model in terms of

$$\begin{aligned} p_{01} = &\quad P(Y_i = 1 | X_i = 0) \\ p_{11} = &\quad P(Y_i = 1 | X_i = 1) \end{aligned}$$

Compute the maximum likelihood estimators $\hat{p}_{01}$ and $\hat{p}_{11}$ for $p_{01}$ and $p_{11}$, respectively. Express your answer in terms of $f_{00}$ (enter "A"), $f_{01}$ (enter "B"), $f_{10}$ (enter "C"), $f_{11}$ (enter "D") and $n$.

$\hat{p}_{01}$ B/(A+B) ✔ **Answer:** B/(A+B)

$\frac{B}{A+B}$

$\hat{p}_{11}$ D/(C+D) ✔ **Answer:** D/(C+D)

$\frac{D}{C+D}$

**Solution:**

The likelihood for the model can be written as

$$\begin{aligned} P(Y_1 = y_1, \ \ldots, Y_n = y_n | X_1 = x_1, \ldots, X_n = x_n) \\ &= \prod_{i=1}^{n} \Big[ p_{01} \mathbf{1}(x_i = 0, y_i = 1) + (1 - p_{01}) \mathbf{1}(x_i = 0, y_i = 0) \\ &\quad + p_{11} \mathbf{1}(x_i = 1, y_i = 1) + (1 - p_{11}) \mathbf{1}(x_i = 1, y_i = 0) \Big] \\ &= p_{01}^{n f_{01}} (1 - p_{01})^{n f_{00}} p_{11}^{n f_{11}} (1 - p_{11})^{n f_{10}}. \end{aligned}$$

Taking logarithms yields

$$\ln P(Y_1 = y_1, \ldots, Y_n = y_n | X_1 = x_1, \ldots, X_n = x_n) = n \left[ f_{01} p_{01} + f_{00} (1 - p_{01}) + f_{11} p_{11} + f_{10} (1 - p_{11}) \right].$$

Differentiating and setting the derivative to zero then leads to the maximum likelihood estimators

$$\begin{aligned} \hat{p}_{01} &= \frac{f_{01}}{f_{01} + f_{00}} \\ \hat{p}_{11} &= \frac{f_{11}}{f_{11} + f_{10}}. \end{aligned}$$

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

## (b)

2/2 points (graded)

Although the $X_i$ are discrete, we can also use a logistic regression model to analyze the data. That is, now we assume

$$Y_i | X_i \sim \text{Ber}\left(\frac{1}{1+\mathbf{e}^{-(X_i\beta_1+\beta_0)}}\right),$$

for $\beta_0, \beta_1 \in \mathbb{R}$, and that given $X_i$, the $Y_i$ are independent.

Calculate the maximum likelihood estimator $\widehat{\beta}_0$, $\widehat{\beta}_1$ for $\beta_0$ and $\beta_1$, where we again assume that all $f_{kl} > 0$. Express your answer in terms of $f_{00}$ (enter "A"), $f_{01}$ (enter "B"), $f_{10}$ (enter "C"), $f_{11}$ (enter "D") and $n$.

$\widehat{\beta}_0$ ln(B/A) ✔ **Answer:** ln(B/A)

$\ln\left(\frac{B}{A}\right)$

$\widehat{\beta}_1$ ln(D/C)-ln(B/A) ✔ **Answer:** ln((A*D)/(B*C))

$\ln\left(\frac{D}{C}\right) - \ln\left(\frac{B}{A}\right)$

**Solution:**

The gradient equations that determines the maximum likelihood estimator the one calculated for logistic regression in class and can be written as

$$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \frac{1}{1+\mathbf{e}^{-x_i\widehat{\beta}_1-\widehat{\beta}_0}} \tag{11.1}$$

$$\sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} x_i \frac{1}{1+\mathbf{e}^{-x_i\widehat{\beta}_1-\widehat{\beta}_0}} \tag{11.2}$$

We note that by counting the elements where $y_i = 1$ and $x_i = 1$,

$$\begin{aligned}
\sum_{i=1}^{n} y_i &= n\left(f_{01}+f_{11}\right) \\
\sum_{i=1}^{n} x_i y_i &= n f_{11} \\
\sum_{i=1}^{n} x_i \frac{1}{1+\mathbf{e}^{-x_i\widehat{\beta}_1-\widehat{\beta}_0}} &= n\left(f_{10}+f_{11}\right)\frac{1}{1+\mathbf{e}^{-\widehat{\beta}_1-\widehat{\beta}_0}} \\
\sum_{i=1}^{n} \frac{1}{1+\mathbf{e}^{-x_i\widehat{\beta}_1-\widehat{\beta}_0}} &= n\left(f_{01}+f_{00}\right)\frac{1}{1+\mathbf{e}^{-\widehat{\beta}_0}} + n\left(f_{10}+f_{11}\right)\frac{1}{1+\mathbf{e}^{-\widehat{\beta}_1-\widehat{\beta}_0}}.
\end{aligned}$$

This means we can rewrite the second gradient equation to

$$f_{11} = \left(f_{10}+f_{11}\right)\frac{1}{1+\mathbf{e}^{-\widehat{\beta}_1-\widehat{\beta}_0}} \iff \mathbf{e}^{-\widehat{\beta}_1-\widehat{\beta}_0} = \frac{f_{10}}{f_{11}}.$$

Plugging this into the first gradient equation then leads to

$$\left(f_{01}+f_{00}\right)\frac{1}{1+\mathbf{e}^{-\widehat{\beta}_0}} + f_1 1 = f_{01}+f_{11} \iff \mathbf{e}^{-\widehat{\beta}_0} = \frac{f_{00}}{f_{01}}.$$

Inserted back into the previous equation, we arrive at

$$\mathbf{e}^{-\widehat{\beta}_1} = \frac{f_{10}f_{01}}{f_{00}f_{11}}.$$

Taking logarithms then finally yields

$$\begin{aligned}\widehat{\beta}_0 = &\ \ln\left(\frac{f_{01}}{f_{00}}\right)\\ \widehat{\beta}_1 = &\ \ln\left(\frac{f_{00}f_{11}}{f_{01}f_{10}}\right).\end{aligned}$$

Submit You have used 3 of 3 attempts

Answers are displayed within the problem

## (c)

2/2 points (graded)

Given the maximum likelihood estimators $\widehat{\beta}_0$ , $\widehat{\beta}_1$ , what are the associated predicted probabilities

$$\begin{aligned}\widetilde{p_{01}} = &\ \ P(Y_i = 1|X_i = 0, \widehat{\beta}_0, \widehat{\beta}_1)\\ \widetilde{p_{11}} = &\ \ P(Y_i = 1|X_i = 1, \widehat{\beta}_0, \widehat{\beta}_1)\end{aligned}$$

in terms of $f_{kl}$, for $k, l \in \{0,1\}$?

Express your answer in terms of $f_{00}$ (enter "A"), $f_{01}$ (enter "B"), $f_{10}$ (enter "C"), $f_{11}$ (enter "D") and $n$.

$\widetilde{p_{01}}$ B/(A+B) ✔ **Answer:** B/(A+B)

$\frac{B}{A+B}$

$\widetilde{p_{11}}$ D/(C+D) ✔ **Answer:** D/(C+D)

$\frac{D}{C+D}$

**Solution:**

We plug the solutions $\widehat{\beta}_0, \widehat{\beta}_1$ back into the associated likelihoods:

$$\begin{aligned} \widetilde{p_{01}} &= \mathbf{P}\left(Y_i = 1 | X_i = 0, \widehat{\beta}_0, \widehat{\beta}_1\right) \\ &= \frac{1}{1+\mathbf{e}^{-\widehat{\beta}_0}} = \frac{1}{1+\frac{f_{00}}{f_{01}}} = \frac{f_{01}}{f_{00}+f_{01}}. \\ \widetilde{p_{11}} &= \mathbf{P}\left(Y_i = 1 | X_i = 1, \widehat{\beta}_0, \widehat{\beta}_1\right) \\ &= \frac{1}{1+\mathbf{e}^{-\widehat{\beta}_0-\widehat{\beta}_1}} = \frac{1}{1+\frac{f_{01}f_{10}f_{00}}{f_{00}f_{11}f_{01}}} = \frac{f_{11}}{f_{10}+f_{11}}. \end{aligned}$$

In fact, this coincides with the result we obtained in (a), so we can conclude that this is merely a re-parametrization of the original Bernoulli model. In this case, the logistic regression model only excludes zeros in the frequencies $f_{kl}$ and otherwise does not pose any restriction.

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

## Discussion

Hide Discussion

**Topic:** Unit 7 Generalized Linear Models:Homework 11 / 4. Contingency tables

Add a Post

Show all posts ▼ by recent activity ▼

[staff]: (b) Beta estimator expression terms 3

[Staff] Typographical Error 2
I think the following log-likelihood has some typo: [mathjax]\displaystyle n \left[ f_{01} p_{01} + f_{00} (1 - p_{01}) + f_{11} p_{11} + f_{10} (1 - p_{11}) \right][/mathjax] It should ...

? [Staff] Why Can't We Review the Midterms Anymore? 1

Looking at the problems of the exams would be a very good way to get prepared for the final. I don't understand why the exam content is no longer available.

Please Review My Answer to b
I did b differently than what is given in the answer. Thought I'd share and see if my way is permissible or if I got lucky.
8

Any hints for b
Been struggling with this for quite some time now. I can find the likelihood and take derivatives w.r.t to both beta_0 and beta_1, but not sure how that would get me an answe...
14

A nice problem to solve, what is the interpretation / insights obtained from the results?
Following
3

Great Lecture on High Level Stuff
Hi Folks, I hope all is well. I just wanted to share a lecture I found that I think really puts what we are learning into perspective. It's more high level stuff but I think it helped me...
3

Part (b): beta_0 and beta_1
3

What does "#" mean?
I am sorry, but after completing the lecture and the exercises I do not understand what the problem is stating: f_00 f_01 etc.
4