Eurisko Assignment 33

Charlie Weinberger

April 7, 2021

1 33-1

Suppose you are given the following dataset:

$$\text{data} = [(1, 0.2), (2, 0.25), (3, 0.5)]$$

Fit a logistic regression model $y = \dfrac{1}{1 + e^{ax+b}}$ **by hand**.

1. Re-express the model in the form $ax + b =$ some function of y (i.e. isolate $ax + b$ in the logistic regression model). Hint: your function of y will involve ln.
2. Set up a system of equations and turn the system into a matrix equation.
3. ind the best approximation to the solution of that matrix equation by using the pseudoinverse.
4. Substitute your solution for the coefficients of the model, and **plot the model along with the 3 given data points on the same graph** to ensure that the model fits the data points well.

Show all of your steps. No code allowed!

Solution

First, we must isolate $ax + b$. We see that

$$\begin{aligned}
\frac{1}{1 + e^{ax+b}} &= y \\
1 + e^{ax+b} &= \frac{1}{y} \\
e^{ax+b} &= \frac{1}{y} - 1 \\
ax + b &= \ln\left(\frac{1}{y} - 1\right)
\end{aligned}$$

Using the data points that we have, we have the equations

$$
\begin{aligned}
a + b &= \ln\left(\frac{1}{0.2} - 1\right) = 1.3863 \\
2a + b &= \ln\left(\frac{1}{0.25} - 1\right) = 1.0986 \\
3a + b &= \ln\left(\frac{1}{0.5} - 1\right) = 0
\end{aligned}
$$

Turning that into a matrix equation, we get

$$
\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 1.3863 \\ 1.0986 \\ 0 \end{bmatrix}
$$

We know that the pseudoinverse of a matrix A is $(A^T A)^{-1} A^T$.

In our case,

$$
A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}
$$

Therefore, the pseudoinverse of A is

$$
\begin{aligned}
(A^T A)^{-1} A^T &= \left(\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \\
&= \frac{1}{3 \cdot 14 - 6 \cdot 6} \begin{bmatrix} 14 & -6 \\ -6 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \\
&= \frac{1}{6} \begin{bmatrix} 8 & 2 & -4 \\ -3 & 0 & 3 \end{bmatrix}
\end{aligned}
$$

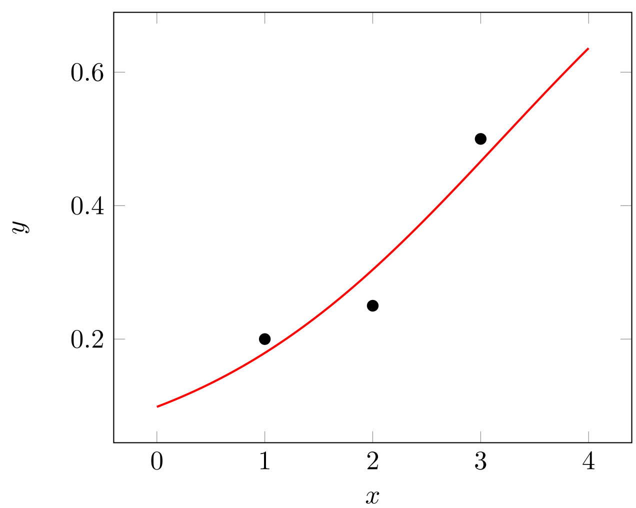
Multiplying both sides by the pseudoinverse, we get

$$\frac{1}{6}\begin{bmatrix} 8 & 2 & -4 \\ -3 & 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} b \\ a \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 8 & 2 & -4 \\ -3 & 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1.3863 \\ 1.0986 \\ 0 \end{bmatrix}$$
$$\frac{1}{6}\begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix} \cdot \begin{bmatrix} b \\ a \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 13.2876 \\ -4.1589 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 2.2146 \\ -0.69315 \end{bmatrix}$$
$$\begin{bmatrix} b & 0 \\ 0 & a \end{bmatrix} = \begin{bmatrix} 2.2146 \\ -0.69315 \end{bmatrix}$$

Therefore,

$$\boxed{y = \frac{1}{1 + e^{(-0.69315x + 2.2146)}}}$$

When we plot the model and out points, we get the following graph:



2 33-2

(a) Given that $X \sim p(x)$, where $p(x)$ is a continuous distribution, prove that for any real number a we have $\mathrm{E}[aX] = a\mathrm{E}[X]$. You should start by writing $\mathrm{E}[aX]$ as an integral, manipulating it, and then simplifying the result into $a\mathrm{E}[X]$. The manipulation will just be 1 step.

We know that

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x \cdot p(x)\,\mathrm{d}x$$

We can use that to show $\mathrm{E}[aX] = a\mathrm{E}[X]$ here:

$$\mathrm{E}[aX] = \int_{-\infty}^{\infty} (ax) \cdot p(x)\,\mathrm{d}x = a \cdot \int_{-\infty}^{\infty} x \cdot p(x)\,\mathrm{d}x = a \cdot (\mathrm{E}[X]) = a\mathrm{E}[X]$$

(b) Given that $X \sim p(x)$ where $p(x)$ is a continuous probability distribution, prove the identity $\mathrm{Var}[X] = \mathrm{E}[X^2] - E[X]^2$.

$$\begin{aligned}
\mathrm{Var}[X] &= \int_{-\infty}^{\infty} (x - E[X])^2 \cdot p(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \left(x^2 - 2x\mathrm{E}[X] + \mathrm{E}[X]^2\right) \cdot p(x)\,\mathrm{d}x \\
&= \left(\int_{-\infty}^{\infty} x^2 \cdot p(x)\,\mathrm{d}x\right) - \left(\int_{-\infty}^{\infty} 2x\mathrm{E}[X] \cdot p(x)\,\mathrm{d}x\right) + \left(\int_{-\infty}^{\infty} \mathrm{E}[X]^2 \cdot p(x)\,\mathrm{d}x\right) \\
&= \mathrm{E}[X^2] - 2\mathrm{E}[X] \cdot \left(\int_{-\infty}^{\infty} x \cdot p(x)\,\mathrm{d}x\right) + \mathrm{E}[X]^2 \cdot \left(\int_{-\infty}^{\infty} p(x)\,\mathrm{d}x\right) \\
&= \mathrm{E}[X^2] - \left(2\mathrm{E}[X] \cdot \mathrm{E}[X]\right) + \left(\mathrm{E}[X]^2 \cdot 1\right) \\
&= \mathrm{E}[X^2] - 2\mathrm{E}[X]^2 + \mathrm{E}[X]^2 \\
\mathrm{Var}[X] &= \mathrm{E}[X^2] - \mathrm{E}[X]^2
\end{aligned}$$