

GroupC Practice

1. For a logistic regression model with $\mathbf{w} = (-1, 2)$ and $b = -5$, find $\hat{P}(y = 1 | \mathbf{x} = (4, 5))$.

(Round to two decimal places.)

ANSWER:

$$\hat{P}(y = 1 | \mathbf{x}) = \frac{1}{1 + e^{-(\mathbf{w}\mathbf{x} + b)}} = \frac{1}{1 + e^{-[(-1, 2) \cdot (4, 5) - 5]}} = \frac{1}{1 + e^{-1}} \approx 0.73$$

2. How do we classify a new point at $\mathbf{x} = (4, 5)$ if using a decision threshold of 0.5?

- ☐ $\hat{y} = 0$
☐ $\hat{y} \approx 0.27$
☐ $\hat{y} \approx 0.73$
☒ $\hat{y} = 1$

3. We adjust logistic regression parameters $\mathbf{w}$ and b to maximize the likelihood of our data according to the model. Or, mathematically equivalently, we minimize the negative log likelihood.

Computationally, the two approaches are not equivalent. One of the reasons we prefer the second approach is that it helps prevent underflow of Python's `float` when calculating products of probabilities. Run

```
import sys
print(f'sys.float_info.min={sys.float_info.min}')
print(f'sys.float_info.max={sys.float_info.max}')
```

to find the `float`'s minimum positive value and maximum value.

- (a) How many probabilities, each around 0.01, can we multiply before underflowing a `float`?

horsepower. ANSWER: $0.01^x = 2.2\text{e-}308 \implies x \approx 153.9 \approx 153$

- (b) What is the negative (natural) log of 0.01?

horsepower. ANSWER: ≈ 4.6

- (c) How many negative log probabilities, each around (b), can we add without overflow?

horsepower. ANSWER: $\approx 1.8\text{e}308 / 4.6 \approx 3.9\text{e}+307$

- (d) Isn't using log likelihood clever?

- ☐ Yes. ANSWER: ☒
☐ It takes more than that to impress me.