

**Logistic Regression and Regularization***(find and work with a partner)*

1. A key step in our derivation of the SGD updates for logistic regression was the fact that  $g'(z) = g(z)(1 - g(z))$ , where  $g(z) = \frac{1}{1+e^{-z}}$ . This allowed us to cancel out the terms in the denominators. Compute the derivative of  $g(z)$  to demonstrate this fact. What does  $g'(z)$  tend to as  $z \rightarrow \infty$ ? As  $z \rightarrow -\infty$ ?
2. The confusion matrices below show hiring predictions separated by demographic group (non-men and men). To put this in the context of our fairness regularization setup, identify  $Y \in \{0, 1\}$  and  $A \in \{0, 1\}$ .

Non-men			Men		
	Predicted don't hire	Predicted DO hire		Predicted don't hire	Predicted DO hire
Test Label wasn't hired	542	170	Test Label wasn't hired	1598	430
Test Label WAS hired	23	56	Test Label WAS hired	340	190

3. Compute the demographic parity:

$$\frac{P(\hat{Y} = 1|A = 1)}{P(\hat{Y} = 1|A = 0)}$$

4. Compute the equalized odds:

$$\frac{P(\hat{Y} = 1|A = 1, Y = y)}{P(\hat{Y} = 1|A = 0, Y = y)} \quad \text{for } y \in \{0, 1\}$$