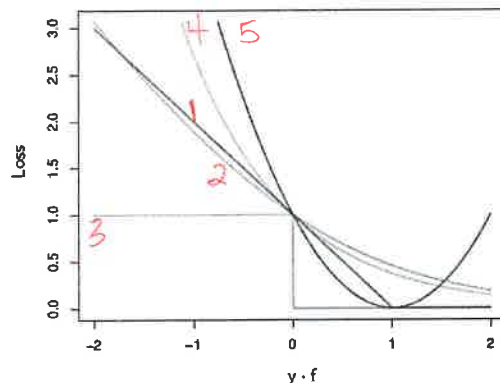


1. Label each of the following functions with its corresponding Loss: 1) Hinge Loss 2) Logistic Loss 3) Zero-One Loss 4) Exponential Loss 5) Squared Loss. (5 points)



2. Explain why you would use any of the above losses in place of the Zero-one loss. (2 points)

These approximations of the zero-one loss allow us to take the derivative and find the minimum of the function. This will give us a  $\vec{w} \cdot \vec{b}$  that have a minimum loss (error).

3. Give an example of a situation where you would want to use the squared loss over the hinge loss. Briefly explain why the squared loss would be better in this case. (3 points)

If you care about being too right or too wrong. An example would be regression. If you're predicting the temperature, you don't want to be way off in either direction.