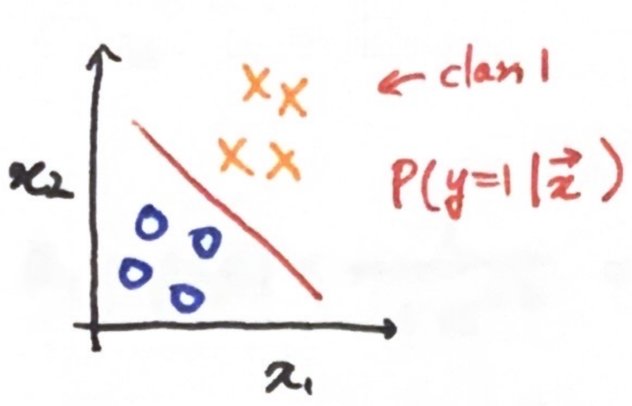
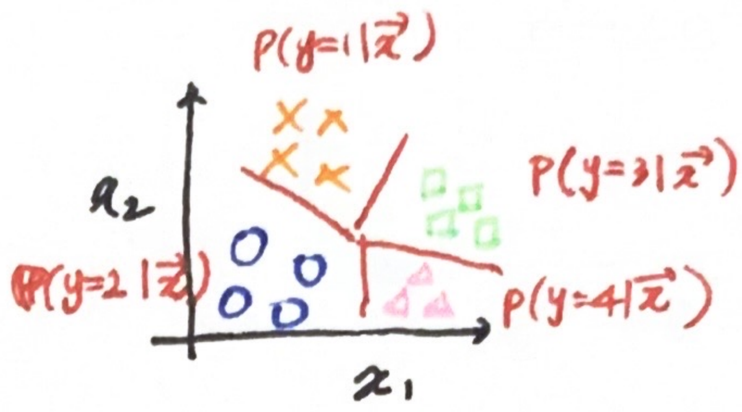
<Multiclass>

- multiclass classification problem: target y can take on more than two possible values

(binary)



(multiclass)



<Softmax>

- Softmax regression algorithm is a generalization of logistic regression which is a binary classification algorithm to the multi-class classification context

* Logistic Regression : 2 possible output values

$$z = \vec{w} \cdot \vec{x} + b$$

$$a_1 = g(z) = \frac{1}{1+e^{-z}} = P(y=1|\vec{x})$$

$$a_2 = 1 - a_1 = P(y=0|\vec{x})$$

$\because P(y=1|\vec{x}) + P(y=0|\vec{x}) = 1$

* Softmax Regression (e.g. 4 possible outputs)

$$z_1 = \vec{w}_1 \cdot \vec{x} + b_1 \Rightarrow a_1 = \frac{e^{z_1}}{e^{z_1}+e^{z_2}+e^{z_3}+e^{z_4}} = P(y=1|\vec{x})$$

$$z_2 = \vec{w}_2 \cdot \vec{x} + b_2 \Rightarrow a_2 = \frac{e^{z_2}}{e^{z_1}+e^{z_2}+e^{z_3}+e^{z_4}} = P(y=2|\vec{x})$$

$$z_3 = \vec{w}_3 \cdot \vec{x} + b_3 \Rightarrow a_3 = \frac{e^{z_3}}{e^{z_1}+e^{z_2}+e^{z_3}+e^{z_4}} = P(y=3|\vec{x})$$

$$z_4 = \vec{w}_4 \cdot \vec{x} + b_4 \Rightarrow a_4 = \frac{e^{z_4}}{e^{z_1}+e^{z_2}+e^{z_3}+e^{z_4}} = P(y=4|\vec{x})$$

generalize ⟹ [Softmax Regression] N possible outputs

$$z_j = \vec{w}_j \cdot \vec{x} + b_j \quad j = 1, \ldots, N$$

$$a_j = \frac{e^{z_j}}{\sum_{k=1}^{N} e^{z_k}} = P(y=j|\vec{x})$$