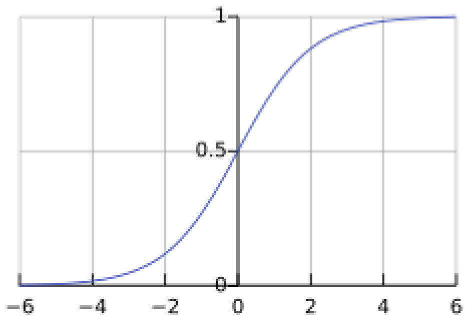
Sigmoid Function

Sigmoid Function is a mathematical function that have "S" shaped curve

$$S(x) = \frac{1}{1+e^{-x}} = \frac{e^x}{e^x+1}$$

- It Convert the X into values that falls in the interval between 1 and 0



- Used in the Logistic Regression which models the **Probability** that a binary outcome variable equals 1 given set of predictors
- $S(0) = 0.5$ That's where the function transitions fastest
- Its Great for modeling **non-linear boundaries**

Formula Explained

$$S(x) = \frac{1}{1+e^{-1}}$$

- e^{-x} inverse of $\frac{1}{e^x}$
- $\lim_{x \to -\infty} e^{-x} = +\infty$
- $\lim_{x \to_0} e^{-x} = 1$
- $\lim_{x \to +\infty} e^{-x} = 0$

Based on the limits of the inverse e^{-x}

$$\lim_{x \to -\infty} S(x) = \frac{1}{1+e^{-1}} = \frac{1}{1+\infty} = 0$$

- The **Sigmoid** goes towards zero when x is negative

$$\lim_{x \to 0} S(x) = \frac{1}{1+e^{-1}} = \frac{1}{1+1} = 0.5$$
$$\lim_{x \to +\infty} S(x) = \frac{1}{1+e^{-1}} = \frac{1}{1+0} = 1$$

- 1 and 0 are the upper and lower bound for the **Sigmoid** function

Sigmoid Function Derivation

The simplest case is the Response Y_i follows a **Bernoulli Distribution**

$$P(Y_i|X_i;\beta) = P(X)^{Y_i}.(1-P(X))^{1-Y_i}$$
$$Y_i \sim \text{Ber}(P(X))$$

With X:

$$\vec{X} = \begin{bmatrix} 1 \\ X_{1,i} \\ X_{2,i} \\ \vdots \\ X_{p,i} \end{bmatrix} \text{ and } \vec{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

- The 1 in the X vector is for the **Intercept**

So we have $P(X)$

$$P(X_i) = \frac{1}{1 + e^{-X_i\beta}}$$

But where did it come from?

We already know that Y_i is Binary, we use **Logit** function also called **log Odds**

$$\operatorname{logit}(\Pr(Y_i = 1)) = \log\left(\frac{\Pr(Y_i = 1)}{\Pr(Y_i = 0)}\right) = \log\left(\frac{P(X)}{1 - P(X)}\right) = \vec{X}\vec{\beta}$$

- Multiply by e on both sides :

$$\frac{P(X)}{1 - P(X)} = e^{\vec{X}\vec{\beta}}$$

$$\begin{aligned} &P(X) = e^{\vec{X}\vec{\beta}}(1 - P(X)) \\ &P(X) = e^{\vec{X}\vec{\beta}} - e^{\vec{X}\vec{\beta}}P(X) \\ &P(X) = \frac{e^{\vec{X}\vec{\beta}}}{1 + e^{\vec{X}\vec{\beta}}} \\ &\text{divied by } e^{\vec{X}\vec{\beta}} \\ &P(X) = \frac{1}{e^{-\vec{X}\vec{\beta}} + 1} \end{aligned}$$