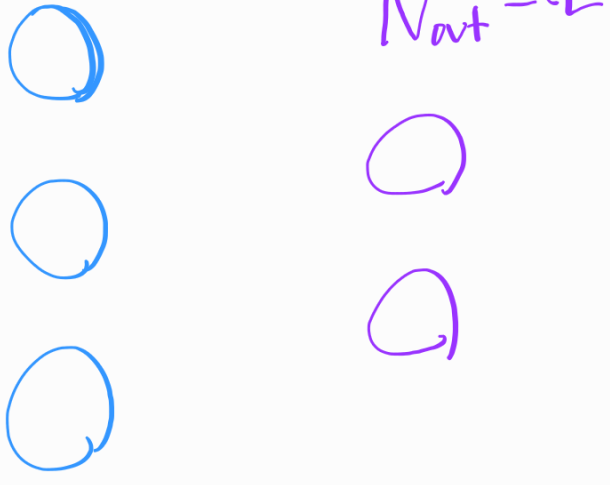
Weight initialization

- **LeCun**
  - $w \sim U\left(-\sqrt{\frac{3}{N_{\text{in}}}}, \sqrt{\frac{3}{N_{\text{in}}}}\right)$ **or** $w \sim N\left(0, \frac{1}{N_{\text{in}}}\right)$
- **Xavier (sigmoid/tanh 사용하는 신경망)**
  - $w \sim U\left(-\sqrt{\frac{6}{N_{\text{in}}+N_{\text{out}}}}, \sqrt{\frac{6}{N_{\text{in}}+N_{\text{out}}}}\right)$ **or** $w \sim N\left(0, \frac{2}{N_{\text{in}}+N_{\text{out}}}\right)$
- **He (ReLU 사용하는 신경망)**
  - $w \sim U\left(-\sqrt{\frac{6}{N_{\text{in}}}}, \sqrt{\frac{6}{N_{\text{in}}}}\right)$ **or** $w \sim N\left(0, \frac{2}{N_{\text{in}}}\right)$

1. 각 Weight는 특정한 Distribution이다.

2. $N_{in}=3$ $N_{out}=2$



3. 평균은 0근처로 잡고 random하게 Sampling.