

k-armed bandits Upper confidence bound

$$\frac{Q_n(a) - a k_n}{m_n k_n}$$

$$R_n$$