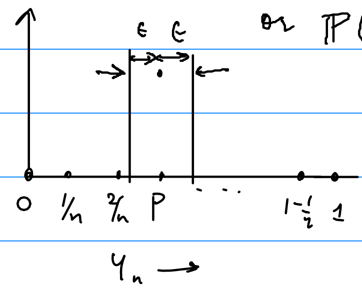
$X_1 \ldots X_n \sim$ independent $B(p)$.

$X_i \in \{0, 1\}$ $\quad P(X_i = 1) = 1/2$

$\mathbb{P}(|Y_n - p| < \epsilon) \geq 1 - ?$

or $\mathbb{P}(|Y_n - p| > \epsilon) \leq ?$



$$Y_n = \frac{X_1 + \ldots + X_n}{n}$$

$$\mathbb{E} Y_n = \mathbb{E}\left(\frac{X_1 + \cdot \quad + X_n}{n}\right)$$

$$= \frac{\mathbb{E}X_1 + \mathbb{E}X_2 + \ldots + \mathbb{E}X_n}{n} = \frac{np}{n} = p.$$

$$\mathbb{E}X_i = 0(1-p) + 1 \cdot p = p.$$

$$\mathbb{E}(Y_n - \mathbb{E}Y_n)^2 = \mathrm{var}(Y_n) = \mathrm{var}\left(\frac{X_1 + \ldots + X_n}{n}\right) = \frac{1}{n^2}\left[\mathrm{var}(X_1 + \ldots + X_n)\right].$$

$$= \frac{1}{n^2}\left[\mathrm{var}(X_1) + \cdots + \mathrm{var}(X_n)\right]$$

$$= \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}.$$

$$\mathrm{var}(X_1) = \mathbb{E}\left[(X_1 - p)^2\right] = (0-p)^2(1-p) + (1-p)^2 p$$

$$= p(1-p)$$

Hoeffding's inequality (Concentration Inequalities).

$X_1 \ldots X_n \sim$ independent random variables $\quad \mathbb{E}X_i = \mu$

$X_i \in [a, b]$. $\quad Y_n = \dfrac{X_1 + \ldots + X_n}{n}$.

$$\mathbb{P}(|Y_n - \mu| > \epsilon) \leq 2\exp\left(-\frac{n\epsilon^2}{2(b-a)^2}\right)$$

$$\mathbb{P}(|Y_n - \mu| < \epsilon) \geq 1 - 2\exp\left(-\frac{n\epsilon^2}{2(b-a)^2}\right)$$

PAC Learning

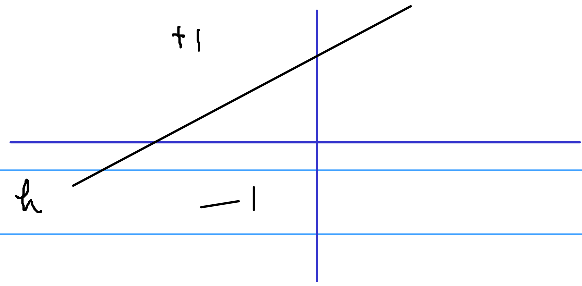
PAC: Probably Approximately Correct

VC dimension

$\mathcal{X}$: instance space. $\quad \mathcal{Y}$: label set $\{-1, 1\}$.

Hypothesis class = Set of "hypothesis".

$\underline{\bar{X}} = \mathbb{R}^2$



$\mathcal{H} = \{\text{all linear classifiers}\}$.

Paul Erdos