❒ **Inégalité d'Hoeffding** – Soit $Z_1, .., Z_m$ $m$ variables iid tirées d'une distribution de Bernoulli de paramètre $\phi$. Soit $\widehat{\phi}$ leur moyenne empirique et $\gamma > 0$ fixé. On a :

$$\boxed{P(|\phi-\widehat{\phi}|>\gamma)\leqslant 2\exp(-2\gamma^2m)}$$

*Remarque : cette inégalité est aussi connue sous le nom de borne de Chernoff.*

❒ **Erreur de training** – Pour un classifieur donné $h$, on définit l'erreur d'entrainement $\widehat{\epsilon}(h)$, aussi connu sous le nom de risque empirique ou d'erreur empirique, par :

$$\boxed{\widehat{\epsilon}(h)=\frac{1}{m}\sum_{i=1}^m 1_{\{h(x^{(i)})\neq y^{(i)}\}}}$$

❒ **Probablement Approximativement Correct (PAC)** – PAC est un cadre dans lequel de nombreux résultats d'apprentissages ont été prouvés, et contient l'ensemble d'hypothèses suivant :

- les jeux d'entrainement et de test suivent la même distribution
- les exemples du jeu d'entrainement sont tirés indépendamment

❒ **Éclatement** – Étant donné un ensemble $S=\{x^{(1)},...,x^{(d)}\}$, et un ensemble de classifieurs $\mathcal{H}$, on dit que $\mathcal{H}$ brise $S$ si pour tout ensemble de labels $\{y^{(1)}, ..., y^{(d)}\}$, on a :

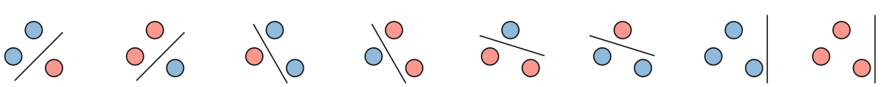
$$\boxed{\exists h\in\mathcal{H}, \quad \forall i\in[\![1,d]\!], \quad h(x^{(i)})=y^{(i)}}$$

❒ **Théorème de la borne supérieure** – Soit $\mathcal{H}$ une hypothèse finie de classe telle que $|\mathcal{H}|=k$, soit $\delta$, et soit $m$ la taille fixée d'un échantillon. Alors, avec une probabilité d'au moins $1-\delta$, on a :

$$\boxed{\epsilon(\widehat{h})\leqslant\left(\min_{h\in\mathcal{H}}\epsilon(h)\right)+2\sqrt{\frac{1}{2m}\log\left(\frac{2k}{\delta}\right)}}$$

❒ **Dimension VC** – La dimension de Vapnik-Chervonenkis (VC) d'une classe d'hypothèses de classes infinies donnée $\mathcal{H}$, que l'on note $\textrm{VC}(\mathcal{H})$, est la taille de l'ensemble le plus grand qui est brisé par $\mathcal{H}$.

*Remarque : la dimension VC de $\mathcal{H}=\{\textrm{set of linear classifiers in 2 dimensions}\}$ est égale à 3.*



❒ **Théorème (Vapnik)** – Soit $\mathcal{H}$ donné, avec $\textrm{VC}(\mathcal{H})=d$ avec $m$ le nombre d'exemples d'entrainement. Avec une probabilité d'au moins $1-\delta$, on a :

$$\boxed{\epsilon(\widehat{h})\leqslant\left(\min_{h\in\mathcal{H}}\epsilon(h)\right)+O\left(\sqrt{\frac{d}{m}\log\left(\frac{m}{d}\right)+\frac{1}{m}\log\left(\frac{1}{\delta}\right)}\right)}$$