Compter jusqu'à n en $O(\log\log n)$ bits

Réf : $\emptyset$

Th : Soit $\varepsilon>0, \delta>0$. Il existe (X_n) une suite de variable aléatoires telles que $\mathbb{P}(|X_n - n| \geq \varepsilon n) < \delta$ pour tout n et presque sûrement, X_n se code en $O(\log\log n)$ bits

Démo • Soit $(\xi_{i,k})_{\substack{i\geq 0\\ k\geq 0}}$ une suite de va de Bernoulli indépendantes telles que $\xi_{i,n} \sim Ber(2^{-i})$
Soit $S_0 = 0$, $S_{n+1} = S_n + \xi_{S_n, n+1}$. Montrons que $2^{S_n}-1$ est proche de n.

• $$\mathbb{E}[2^{S_{n+1}}-1] = \sum_{k\geq 0} \mathbb{E}[(2^{S_{n+1}}-1)\mathbb{1}_{S_n=k}] \quad \text{par cv monotone}$$
$$= \sum_{k\geq 0} \left(\mathbb{E}[2^k . 2^{\xi_{k,n+1}} \mathbb{1}_{S_n=k}] - \mathbb{E}[\mathbb{1}_{S_n=k}]\right)$$
$$= \sum_{k\geq 0} (2^k \mathbb{E}[2^{\xi_{k,n+1}}] - 1) \mathbb{E}[\mathbb{1}_{S_n=k}] \quad \text{par indép}$$
$$= \sum_{k\geq 0} \Big(\underbrace{2^k(2^1 . 2^{-k} + \overset{2^0}{1} . (1-2^{-k}) - 1)}_{=2^k}\Big) \mathbb{E}[\mathbb{1}_{S_n=k}]$$
$$= \sum_{k\geq 0} \mathbb{E}[2^k \mathbb{1}_{S_n=k}] = \mathbb{E}[2^{S_n}]$$

Ainsi, on a $\mathbb{E}[(2^{S_{n+1}}-1)-(n+1)] = \mathbb{E}[(2^{S_n}-1)-n] = \mathbb{E}[(2^{S_0}-1)-0] = 0$
D'où $\mathbb{E}[2^{S_n}-1] = n$

• $$Var(2^{S_n}-1) = Var(2^{S_n}) = \mathbb{E}[(2^{S_n})^2] - \mathbb{E}[2^{S_n}]^2$$
$$= \mathbb{E}[4^{S_n}] - (n+1)^2$$

$$\mathbb{E}[4^{S_{n+1}} - (n+2)^2] = \sum_{k\geq 0} \mathbb{E}[(4^{S_{n+1}} - (n+2)^2)\mathbb{1}_{S_n=k}] \quad \text{par cv monotone}$$
$$= \sum_{k\geq 0} \mathbb{E}[(4^k . 4^{\xi_{k,n+1}} - (n+1)^2 - 2(n+1) - 1)\mathbb{1}_{S_n=k}]$$
$$\perp$$
$$= \sum_{k\geq 0} \Big[\underbrace{4^k . (4^1 . 2^{-k} + 4^0(1-2^{-k}))}_{=4^k+3.2^k} - (n+1)^2 - 2(n+1) - 1\Big] \mathbb{E}[\mathbb{1}_{S_n=k}]$$

$$= \sum_{k\geq 0} \left(\mathbb{E}[(4^{S_n} - (n+1)^2)\mathbb{1}_{S_n=k}] + 3\,\mathbb{E}[2^{S_n}\mathbb{1}_{S_n=k}] - (n+1)^2 - 2(n+1) - 1\right)\mathbb{E}(\mathbb{1}_{S_n=k})$$

$$= \mathbb{E}[4^{S_n} - (n+1)^2] + 3(\mathbb{E}[2^{S_n}-1]+1) - 2(n+1) - 1$$
$$= \mathbb{E}[4^{S_n}-(n+1)^2] + 3n + 3 - 2(n+1) - 1 = \mathbb{E}[4^{S_n}-(n+1)^2] + n$$

D'où $\mathrm{Var}(2^{S_n}-1) = \dfrac{n(n-1)}{2}$

- On prends maintenant k copies indep de X_n : $X_n^{(1)}, \dots, X_n^{(k)}$ et on pose $\bar{X}_n = \frac{1}{k}\sum X_n^{(i)}$

On a $\mathbb{E}(\bar{X}_n) = \mathbb{E}[X_n^{(1)}] = n$

et $\mathrm{Var}(\bar{X}_n) \underset{\perp}{=} \dfrac{1}{k^2}\sum \mathrm{Var}(X_n^{(i)}) = \dfrac{1}{k}\dfrac{n(n+1)}{2}$

Ainsi, $\mathbb{P}(|\bar{X}_n - n| \geq \varepsilon n) \leq \dfrac{\mathrm{Var}(\bar{X}_n)}{(\varepsilon n)^2} \leq \dfrac{1}{2k\varepsilon^2} \leq \delta$ pour k assez grand.

- Montrons un résultat qualitatif sur l'ordre de grandeur des $S_n^{(k)}$:

presque sûrement $N_n := (\lfloor \log_2 S_n^{(1)}\rfloor + 1) + \dots + (\lfloor \log_2 S_n^{(k)}\rfloor+1) < 3k \log_2\log_2 n$.

$$\mathbb{P}[N_n \geq 3k\log_2\log_2 n] \leq \mathbb{P}[\lfloor\log_2 S_n^{(1)}\rfloor + 1 \geq 3\log_2\log_2 n \text{ ou } \dots \text{ ou } \lfloor \log_2 S_n^{(k)}\rfloor+1 \geq 3\log_2\log_2 n]$$
$$\leq k\,\mathbb{P}[\lfloor \log_2 S_n\rfloor + 1 \geq 3\log_2\log_2 n]$$
$$\leq k\,\mathbb{P}[\log_2 S_n + 1 \geq 3 \log_2\log_2 n]$$
$$\leq k\,\mathbb{P}[\log_2 S_n \geq 2\log_2\log_2 n] \quad (\text{pour } n \geq 4)$$
$$\leq k\,\mathbb{P}[S_n \geq (\log_2 n)^2]$$
$$\leq k\,\mathbb{P}[2^{S_n}-1 \geq n^{\log_2 n} - 1]$$
$$\leq k\, \frac{\mathrm{Var}(2^{S_n}-1)}{n^{\ln n} - 1} \leq k\cdot\frac{n^2}{2}\,\frac{1}{n^4} \leq \frac{k}{2}\cdot\frac{1}{n^2}$$

Ainsi $\sum_{n\geq 0} \mathbb{P}(N_n \geq 3k\log_2\log_2 n) < +\infty$ et $\underset{\text{BC}}{\Longrightarrow}$ $\mathbb{P}(\exists N, \forall n \geq N, N_n < 3k\log_2\log_2 n) = 1$