

Informatique et aléatoire: convergence

Nicolas Gast

23 mars 2015

Définir $\lim_{n \rightarrow \infty} X_n$: pourquoi ?

- ▶ Modèles de performances (ex : réseau, calcul)
- ▶ Modèles de comportement (ex : réseaux sociaux)
- ▶ Algorithmes randomisés, algorithmique distribuée
- ▶ Algorithmique sur les grands graphes (e.g., codage)

Exemple : urnes de Polya (voir sujet popularité)

$$X_1 = 1, Y_1 = 1.$$

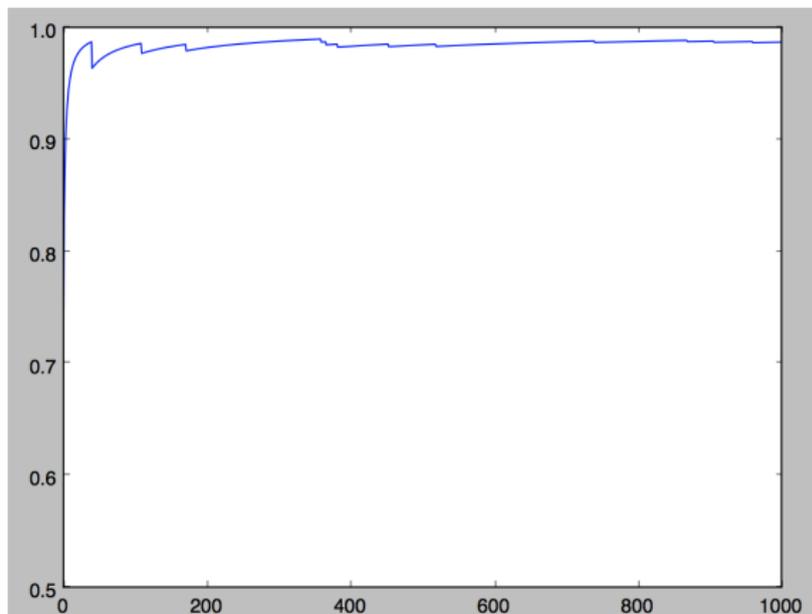
- ▶ Avec probabilité $X_n/(X_n + Y_n)$: $X_{n+1} = X_n + 1$
- ▶ Avec probabilité $Y_n/(X_n + Y_n)$: $Y_{n+1} = Y_n + 1$

Exemple : urnes de Polya (voir sujet popularité)

$$X_1 = 1, Y_1 = 1.$$

- ▶ Avec probabilité $X_n/(X_n + Y_n)$: $X_{n+1} = X_n + 1$
- ▶ Avec probabilité $Y_n/(X_n + Y_n)$: $Y_{n+1} = Y_n + 1$

$$X_n/(X_n + Y_n)$$

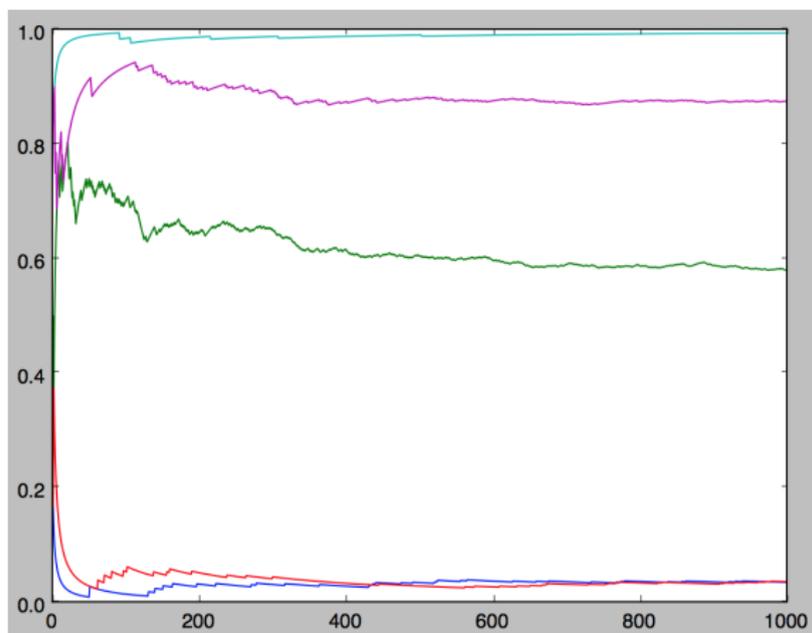


Exemple : urnes de Polya (voir sujet popularité)

$X_1 = 1, Y_1 = 1.$

- ▶ Avec probabilité $X_n/(X_n + Y_n)$: $X_{n+1} = X_n + 1$
- ▶ Avec probabilité $Y_n/(X_n + Y_n)$: $Y_{n+1} = Y_n + 1$

$X_n/(X_n + Y_n)$

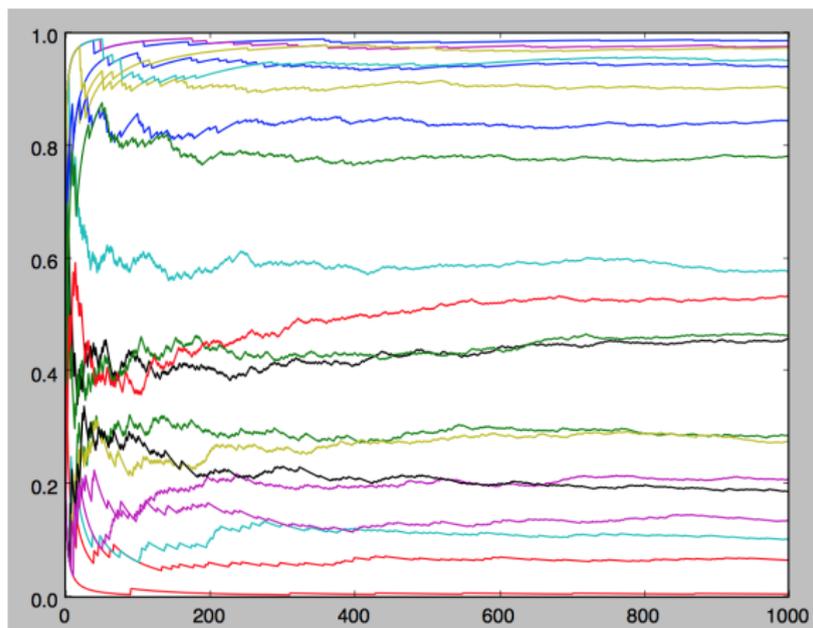


Exemple : urnes de Polya (voir sujet popularité)

$X_1 = 1, Y_1 = 1.$

- ▶ Avec probabilité $X_n/(X_n + Y_n)$: $X_{n+1} = X_n + 1$
- ▶ Avec probabilité $Y_n/(X_n + Y_n)$: $Y_{n+1} = Y_n + 1$

$X_n/(X_n + Y_n)$



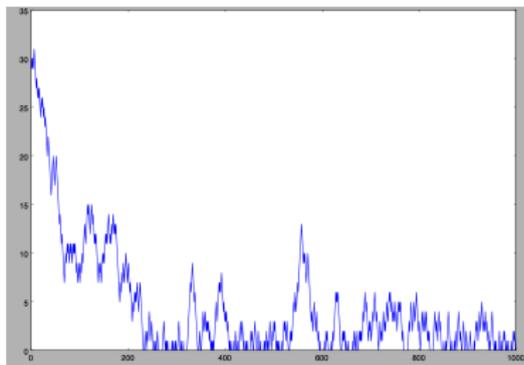
Modèle de réseaux (voir sujet “page rank”)

$$X_{n+1} = \max(X_n + Y_n, 0) \quad \mathbb{P}(Y = 1) = p, \mathbb{P}(Y = -1) = 1 - p.$$

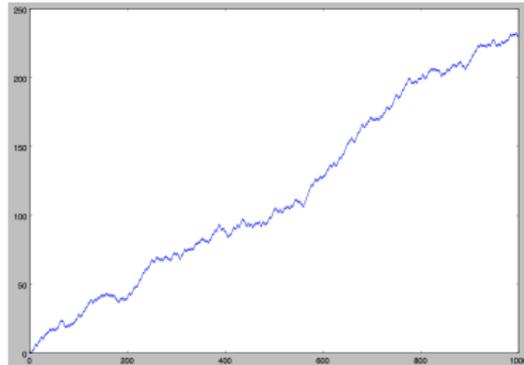
Modèle de réseaux (voir sujet "page rank")

$$X_{n+1} = \max(X_n + Y_n, 0) \quad \mathbb{P}(Y = 1) = p, \mathbb{P}(Y = -1) = 1 - p.$$

$p = 0.4$



$p = 0.6$

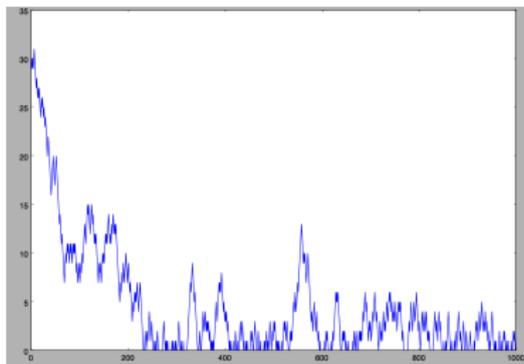


1 traj

Modèle de réseaux (voir sujet “page rank”)

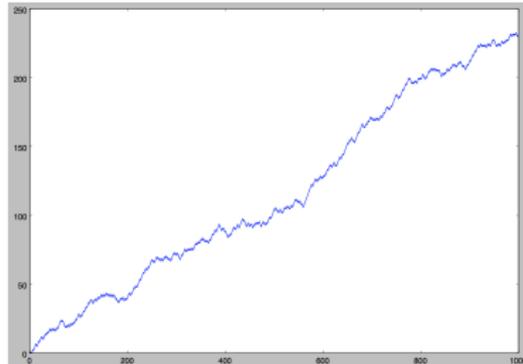
$$X_{n+1} = \max(X_n + Y_n, 0) \quad \mathbb{P}(Y = 1) = p, \mathbb{P}(Y = -1) = 1 - p.$$

$p = 0.4$

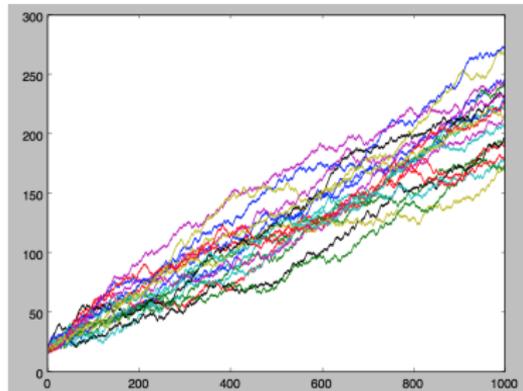
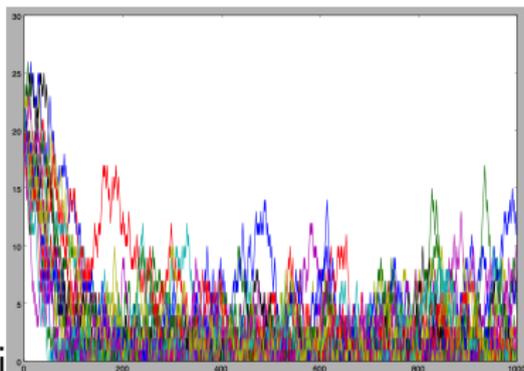


1 traj

$p = 0.6$



20 traj



Quiz

$$\mathbb{P}(X_n = 1) = 1 - \mathbb{P}(X_n = 0) = \frac{1}{2} .$$

Deux propositions :

1. X_n converge ?
2. $\frac{1}{n} \sum X_n$ converge ?

- (A) 1 est vraie, 2 est fausse
- (B) 2 est vraie, 1 est fausse
- (C) 1 et 2 sont vraies
- (D) 1 et 2 sont fausses
- (E) Je ne sais pas.

Quiz (2)

$$\mathbb{P}(X_n = 1) = 1 - \mathbb{P}(X_n = 0) = \frac{1}{2} - \frac{1}{n+1}.$$

Deux propositions :

1. X_n converge ?
 2. $\frac{1}{n} \sum X_n$ converge ?
- (A) 1 est vraie, 2 est fausse
(B) 2 est vraie, 1 est fausse
(C) 1 et 2 sont vraies
(D) 1 et 2 sont fausses
(E) Je ne sais pas.

Quiz (3)

$$\mathbb{P}(Z = 1) = \mathbb{P}(Z = 0) = \frac{1}{2}.$$

$$\blacktriangleright X_n = \begin{cases} Z & \text{si } n = 1 \\ 1 - X_{n-1} & \text{sinon} \end{cases}$$

- (A) X converge
- (B) X ne converge pas
- (C) Je ne sais pas.

Solution : il existe plusieurs notions de convergence

Convergence presque sûre

Définition et exemples

Loi des grands nombres

Modèles de population

Résumé de l'épisode (précédent)

Convergence en loi

Définition et exemples

Théorème central limite et intervalles de confiance

A retenir

Solution : il existe plusieurs notions de convergence

Convergence presque sûre

Définition et exemples

Loi des grands nombres

Modèles de population

Résumé de l'épisode (précédent)

Convergence en loi

Définition et exemples

Théorème central limite et intervalles de confiance

A retenir

Quiz

Que veut dire “A est vrai presque sûrement” :

- (A) A est toujours vrai
- (B) $\mathbb{P}(A) \approx 1$ (A est vrai avec grande probabilité)
- (C) $\mathbb{P}(A) = 1$

Convergence presque sûre (et en probabilité)

(au tableau)

Exercice / Quiz : Ruine d'un joueur

On joue à pile ou face. Si au coup n , un joueur mise y , il gagne y avec probabilité $1/2$ et perd y avec probabilité $1/2$.

La fortune initiale est $X_0 = 0$ et la mise initiale est $y_0 = 1$.

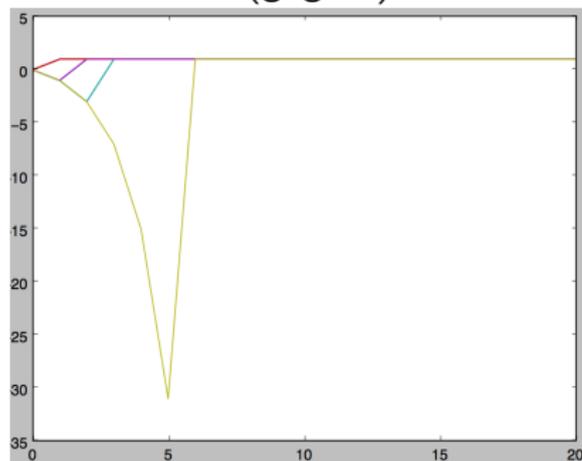
- ▶ Il double la mise tant qu'il n'a pas gagné.

On note X_n sa fortune au temps n .

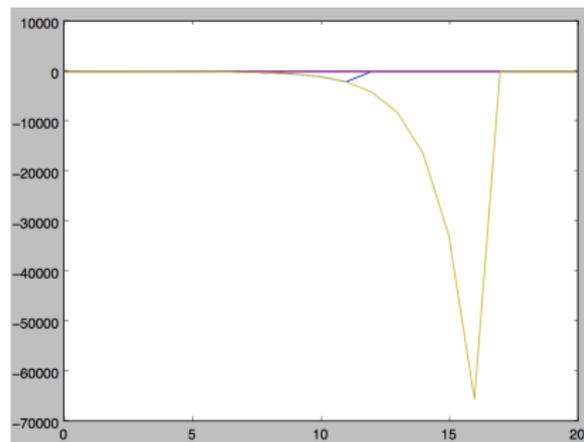
- (A) X_n converge vers 0 presque sûrement.
- (B) X_n converge vers 1 presque sûrement.
- (C) X_n ne converge pas.
- (D) Je ne sais pas.

Ruine : simulation

Proba(gagner)=.5



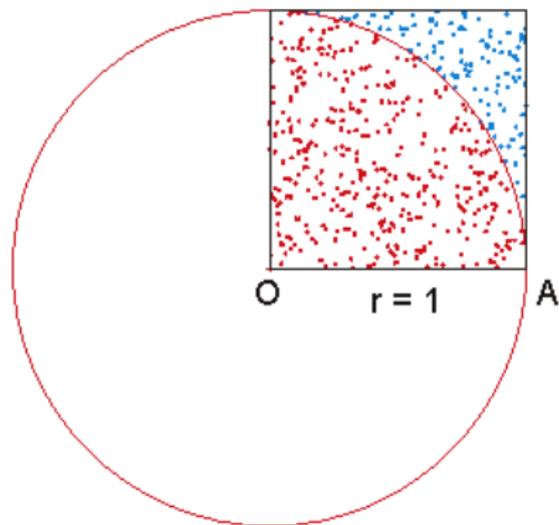
Proba(gagner)=.1



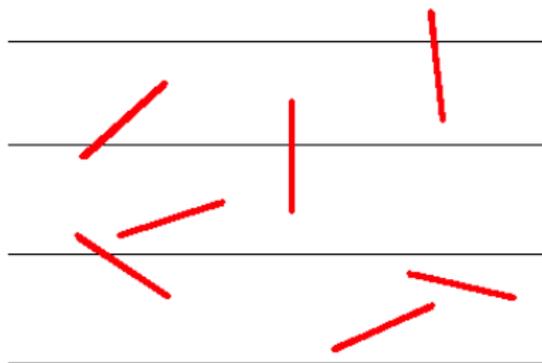
Loi faible / forte des grands nombres

(au tableau)

Application : calcul de π



Application : calcul de π



Exercice (TD puis TP) : processus de branchement

Y_i^n est une suite de variables *i.i.d.*

La population X_n est définie par :

$$X_0 = 1$$

$$X_{n+1} = \sum_{i=1}^{X_n} Y_i^n.$$

Exercice :

1. Calculer $\mathbb{E}(X_n)$ en fonction de $\mathbb{E}(Y)$
2. En déduire que :
 - ▶ si $\mathbb{E}(Y_n) < 1$, alors la population s'éteint presque sûrement.
 - ▶ si $\mathbb{E}(Y_n) > 1$, alors la population converge vers l'infini avec probabilité positive.

Quiz : vrai/faux

$$\mathbb{P}(Z = 1) = \mathbb{P}(Z = 0) = \frac{1}{2}.$$

$$\blacktriangleright X_n = \begin{cases} Z & \text{si } n = 1 \\ 1 - X_{n-1} & \text{sinon} \end{cases}$$

- (A) X_n converge presque sûrement ou en probabilité.
- (B) X_n ne converge pas (presque sûrement ou en probabilité).

Quiz : vrai/faux

$$\mathbb{P}(Y_n = 1) = \mathbb{P}(Y_n = 0) = \frac{1}{2}.$$

$$\blacktriangleright X_n = \sum_{k=1}^n \frac{1}{k} Y_k$$

(A) X_n converge presque sûrement ou en probabilité.

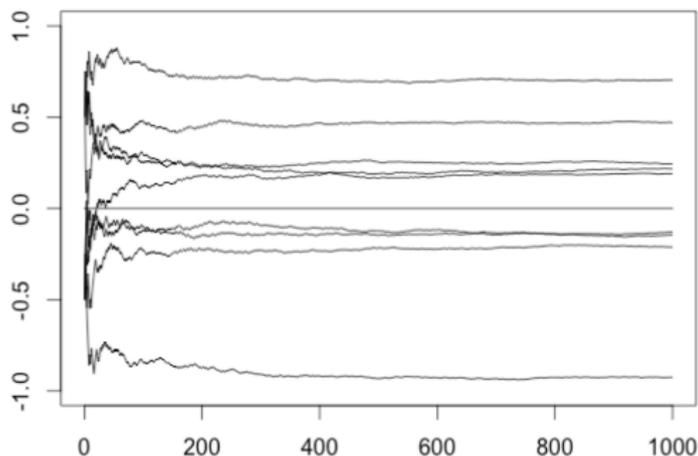
(B) X_n ne converge pas (presque sûrement ou en probabilité).

Quiz : vrai/faux

$$\mathbb{P}(Y_n = 1) = \mathbb{P}(Y_n = 0) = \frac{1}{2}.$$

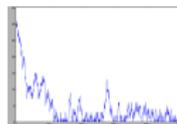
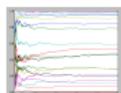
$$\blacktriangleright X_n = \sum_{k=1}^n \frac{1}{k} Y_k$$

- (A) X_n converge presque sûrement ou en probabilité.
- (B) X_n ne converge pas (presque sûrement ou en probabilité).



Résumé de l'épisode

Deux types de convergence.



Convergence presque sûre
Loi des grands nombre



Convergence presque sûre

Définition et exemples

Loi des grands nombres

Modèles de population

Résumé de l'épisode (précédent)

Convergence en loi

Définition et exemples

Théorème central limite et intervalles de confiance

A retenir

Plan du cours

Convergence presque sûre

Définition et exemples

Loi des grands nombres

Modèles de population

Résumé de l'épisode (précédent)

Convergence en loi

Définition et exemples

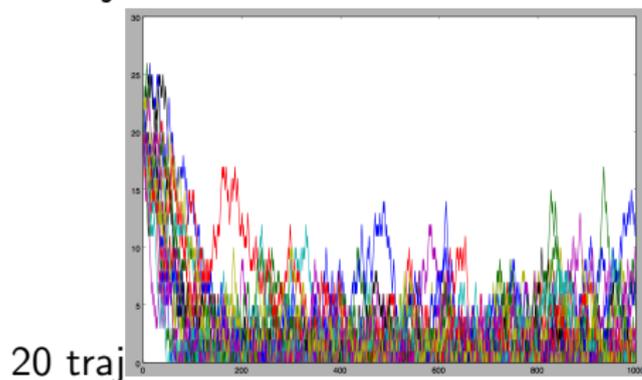
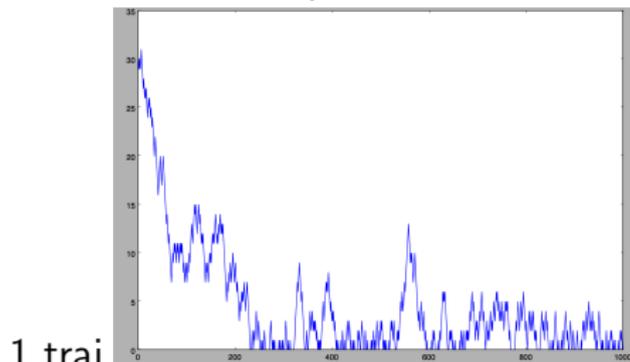
Théorème central limite et intervalles de confiance

A retenir

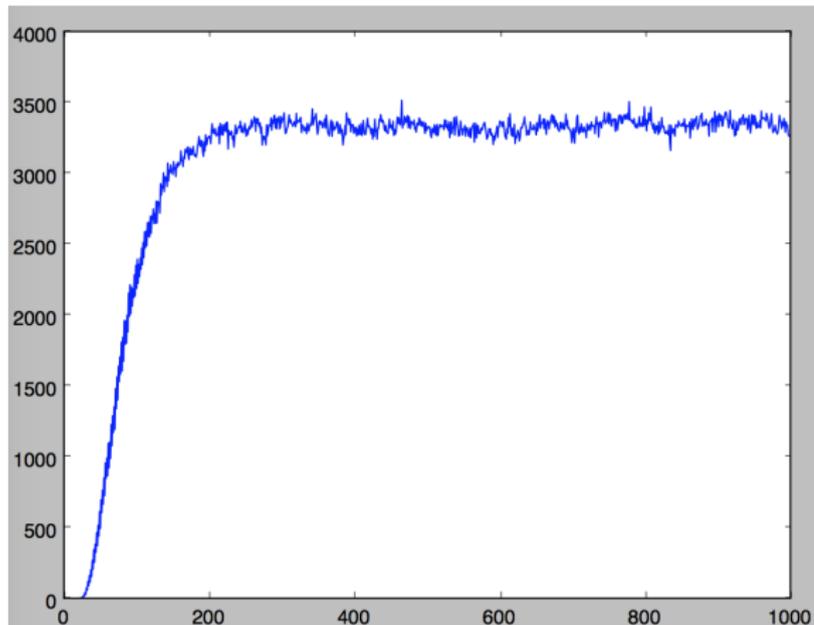
Modèle de réseaux : rappel

$$X_{n+1} = \max(X_n + Y_n, 0) \quad \mathbb{P}(Y = 1) = p, \mathbb{P}(Y = -1) = 1 - p.$$

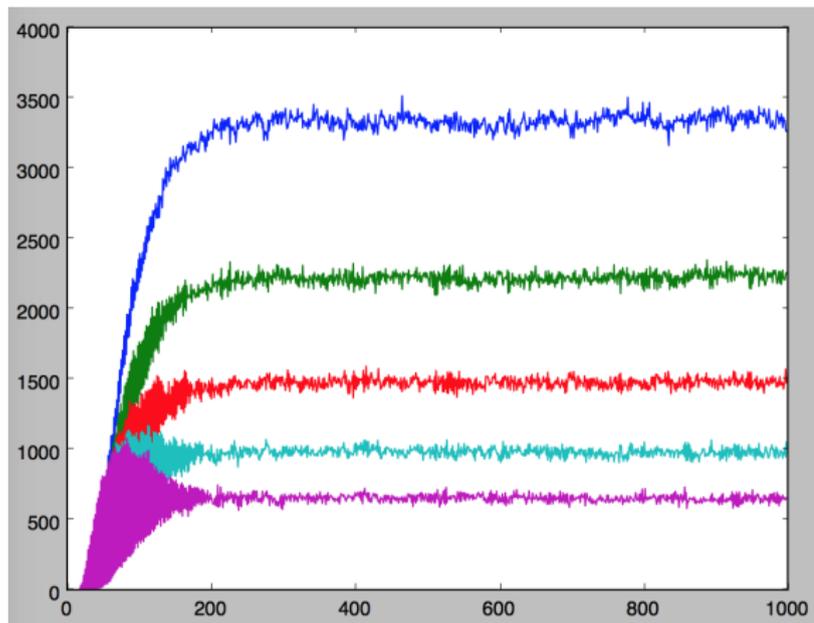
$$p = 0.4$$



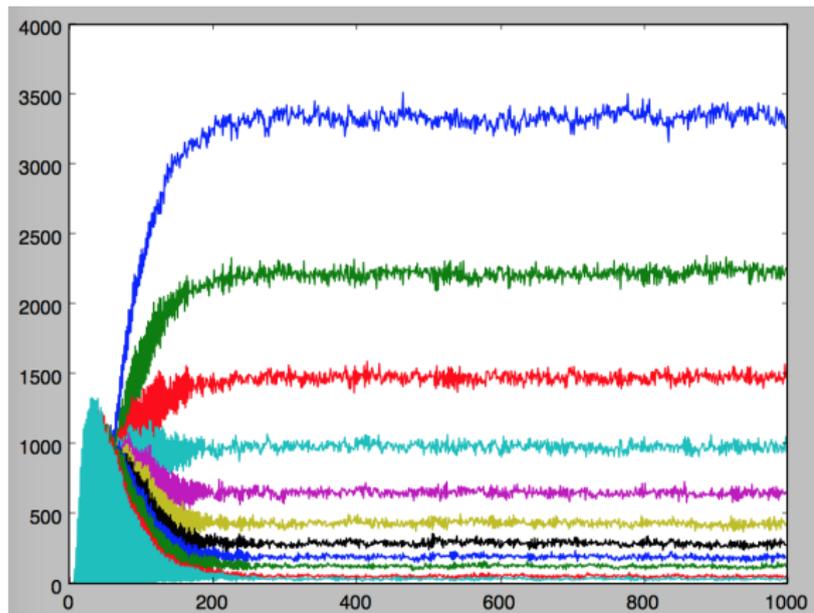
Modèle de réseaux : fréquences empiriques d'apparition



Modèle de réseaux : fréquences empiriques d'apparition



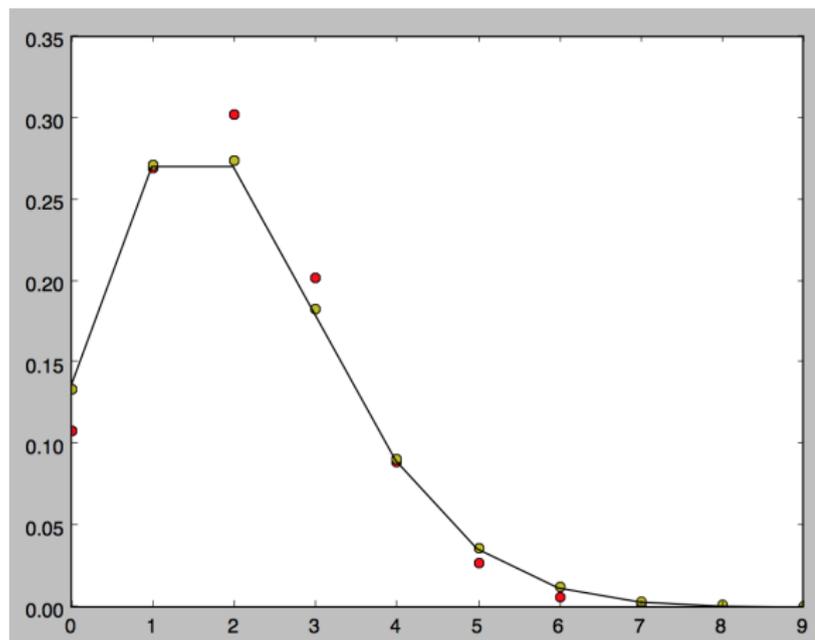
Modèle de réseaux : fréquences empiriques d'apparition



Convergence en loi, exemple de la loi de Poisson

(au tableau)

Convergence de la loi binomiale



Exercice : Poisson

Soit $X \sim \text{Poisson}(\lambda)$ et $Y \sim \text{Poisson}(\mu)$ deux variables indépendantes.

$X + Y$ est :

- (A) une loi de Poisson de paramètre $\lambda\mu$;
- (B) une loi de Poisson de paramètre $\lambda + \mu$;
- (C) une loi de déterministe de paramètre $\lambda\mu$;
- (D) je ne sais pas.

Confiance ?

On dispose d'une pièce biaisée $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p$.

- ▶ On cherche à estimer p .

On effectue des tirages : 0,

Confiance ?

On dispose d'une pièce biaisée $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p$.

- ▶ On cherche à estimer p .

On effectue des tirages : 0, 1, 0, 1, 0, 1,

Confiance ?

On dispose d'une pièce biaisée $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p$.

► On cherche à estimer p .

On effectue des tirages : 0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1,
1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1,
1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0,
0, 1, 1, 0, 1, 0, 0, 0,

Confiance ?

On dispose d'une pièce biaisée $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p$.

► On cherche à estimer p .

On effectue des tirages : 0, 1, 0, 1, 0, 1,1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1,
1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1,
1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0,
0, 1, 1, 0, 1, 0, 0, 0,0, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0,
1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1,
1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1,
0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1,...

Confiance ?

On dispose d'une pièce biaisée $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p$.

► On cherche à estimer p .

On effectue des tirages : 0, 1, 0, 1, 0, 1,1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1,
1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1,
1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0,
0, 1, 1, 0, 1, 0, 0, 0,0, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0,
1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1,
1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1,
0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1,...

	1						
$\frac{1}{n} \sum_{i=1}^n X_i$	0.0						

Confiance ?

On dispose d'une pièce biaisée $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p$.

► On cherche à estimer p .

On effectue des tirages : 0, 1, 0, 1, 0, 1,1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1,
1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1,
1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0,
0, 1, 1, 0, 1, 0, 0, 0,0, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0,
1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1,
1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1,
0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1,...

	1	3	5				
$\frac{1}{n} \sum_{i=1}^n X_i$	0.0	0.333...	.4				

Confiance ?

On dispose d'une pièce biaisée $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p$.

► On cherche à estimer p .

On effectue des tirages : 0, 1, 0, 1, 0, 1,1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1,
1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1,
1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0,
0, 1, 1, 0, 1, 0, 0, 0,0, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0,
1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1,
1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1,
0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1,...

	1	3	5	10	20	100	
$\frac{1}{n} \sum_{i=1}^n X_i$	0.0	0.333...	.4	0.7	0.7	0.57	

Confiance ?

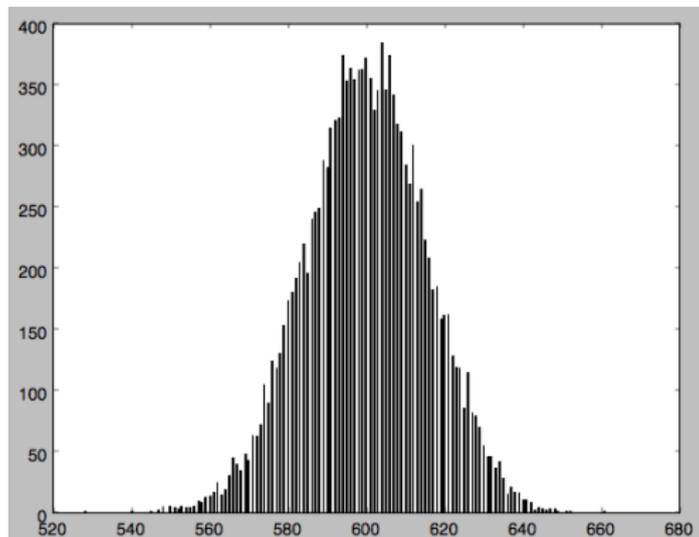
On dispose d'une pièce biaisée $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p$.

► On cherche à estimer p .

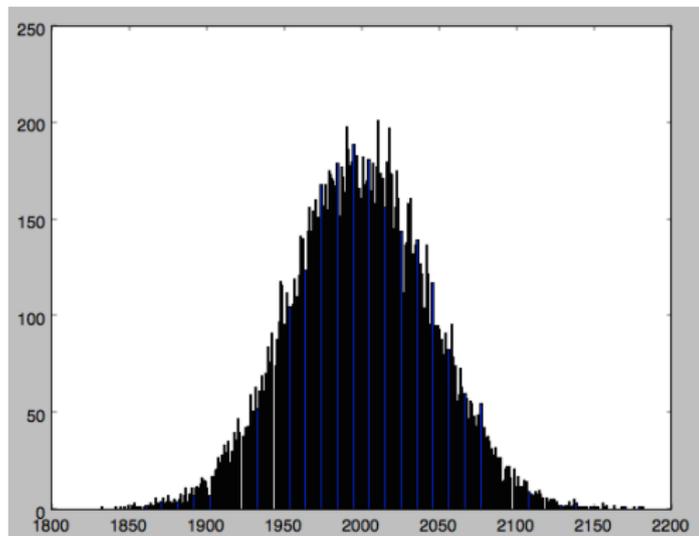
On effectue des tirage : 0, 1, 0, 1, 0, 1,1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1,
1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1,
1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0,
0, 1, 1, 0, 1, 0, 0, 0,0, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0,
1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1,
1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1,
0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1,...

	1	3	5	10	20	100	1000
$\frac{1}{n} \sum_{i=1}^n X_i$	0.0	0.333...	.4	0.7	0.7	0.57	0.606

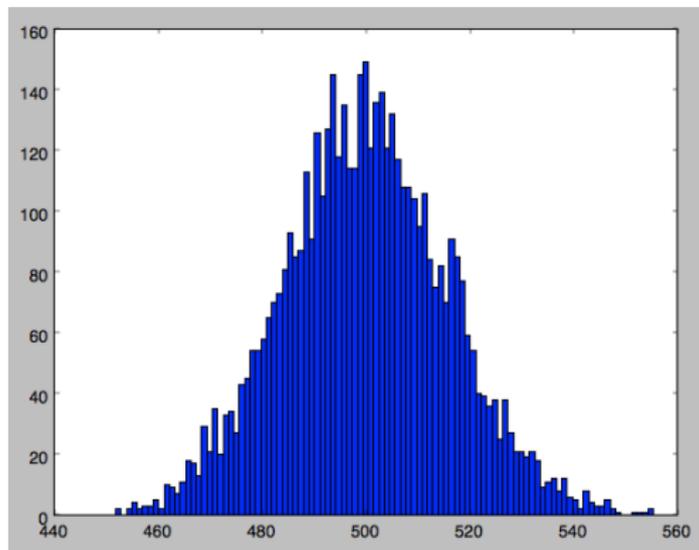
Histogramme de $\sum_{i=1}^{1000} X_i$



Avec des géométriques : $\mathbb{P}(X_i = j) = p(1 - p)^j$.



Avec une exponentielle : $\mathbb{P}(X_i \leq x) = \exp(-\lambda x)$



Théorème central limite : exemple de la marche aléatoire

$$\mathbb{P}(X_n = 1) = \mathbb{P}(X_n = -1) = 0.5. \quad S_n = \sum_{k=1}^n X_k.$$

- ▶ A quelle vitesse converge $\frac{1}{n}S_n$ vers 0 ?

Théorème central limite : exemple de la marche aléatoire

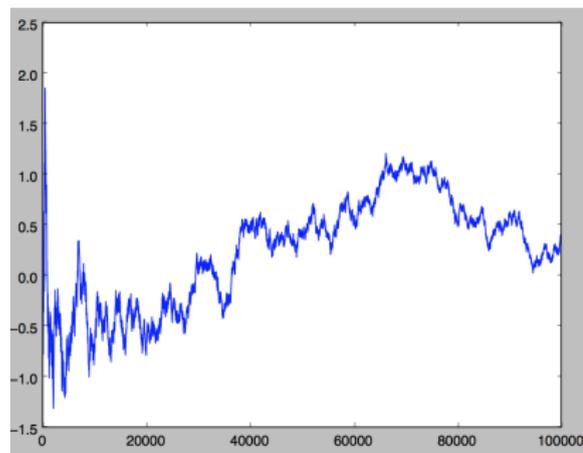
$$\mathbb{P}(X_n = 1) = \mathbb{P}(X_n = -1) = 0.5. \quad S_n = \sum_{k=1}^n X_k.$$

- ▶ A quelle vitesse converge $\frac{1}{n}S_n$ vers 0 ?
- ▶ Est-ce que $\frac{1}{\sqrt{n}}S_n$ converge ?

Théorème central limite : exemple de la marche aléatoire

$$\mathbb{P}(X_n = 1) = \mathbb{P}(X_n = -1) = 0.5. \quad S_n = \sum_{k=1}^n X_k.$$

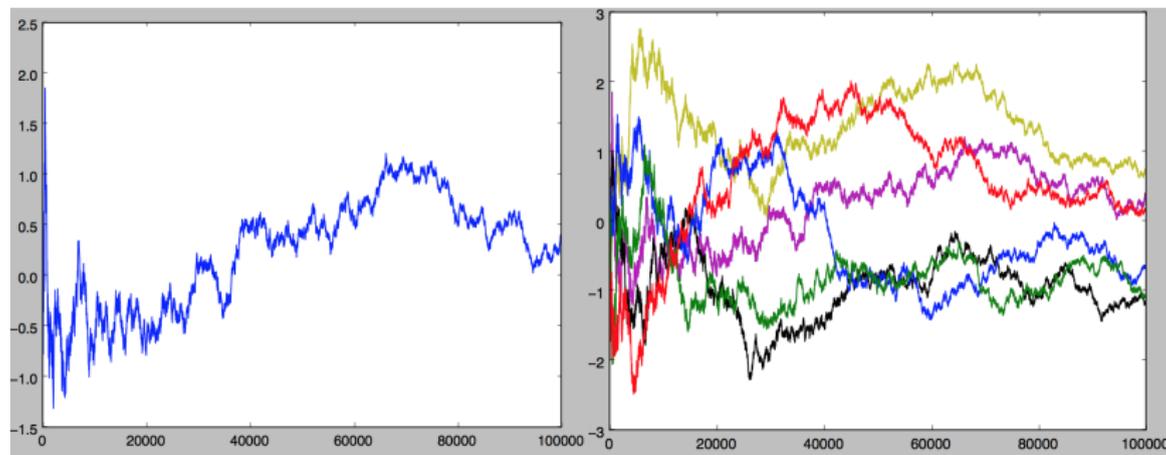
- ▶ A quelle vitesse converge $\frac{1}{n}S_n$ vers 0 ?
- ▶ Est-ce que $\frac{1}{\sqrt{n}}S_n$ converge ?



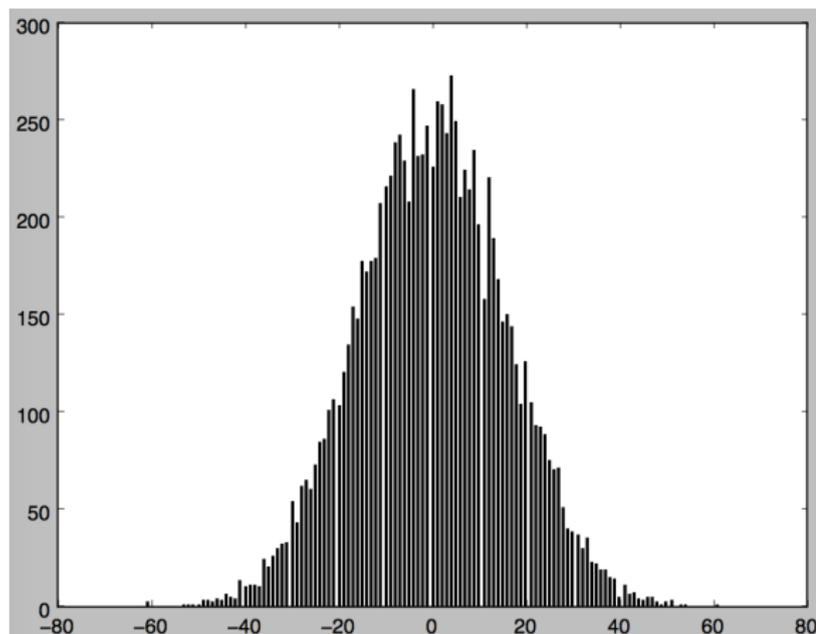
Théorème central limite : exemple de la marche aléatoire

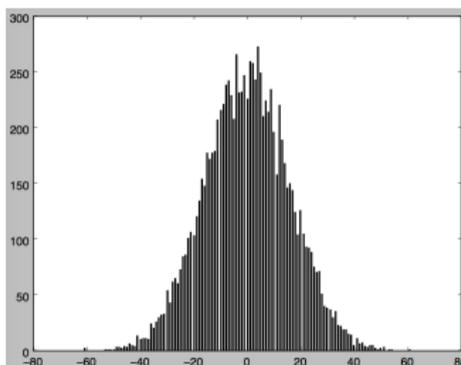
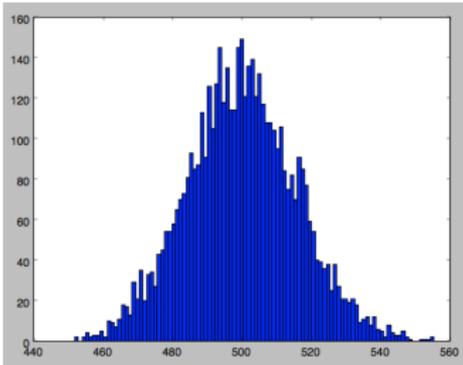
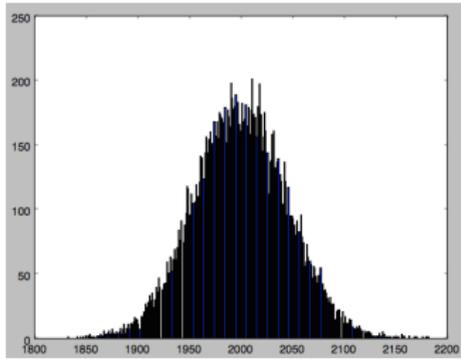
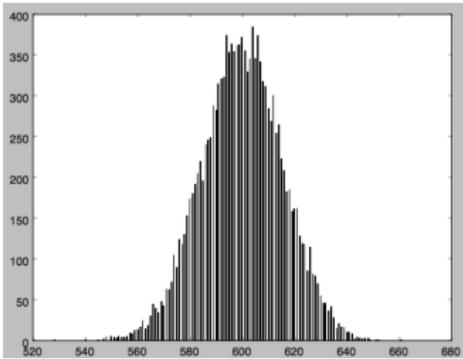
$$\mathbb{P}(X_n = 1) = \mathbb{P}(X_n = -1) = 0.5. \quad S_n = \sum_{k=1}^n X_k.$$

- ▶ A quelle vitesse converge $\frac{1}{n}S_n$ vers 0 ?
- ▶ Est-ce que $\frac{1}{\sqrt{n}}S_n$ converge ?



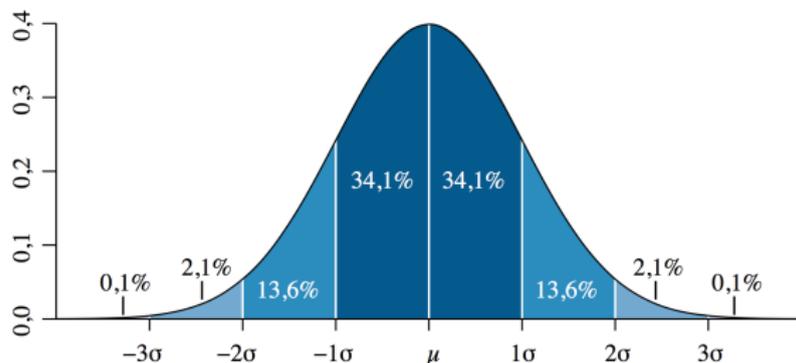
Histogramme des S_{1000} : les valeurs sont (principalement) entre -40 et 40.





Loi normale

$$\text{Densité : } f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$



Théorème central limite et intervalles de confiance sur la moyenne

(au tableau)

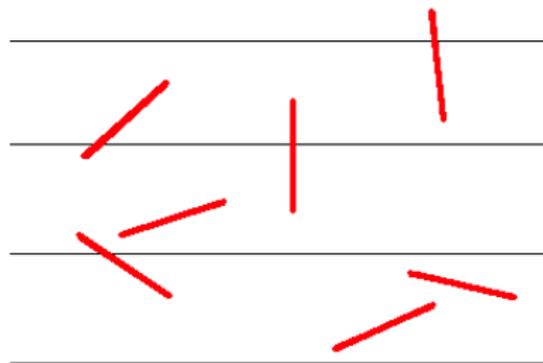
Exercice : application aux sondages

Question X_k : $\mathbb{P}(X_k = 1) = 1 - \mathbb{P}(X_k = 0) = p$.

$$S_n = \sum_{k=1}^n X_k$$

1. Calculer la variance de X_k
2. En déduire la variance de S_n/n
3. Pourquoi dit-on souvent que la marge d'erreur d'un sondage est de 3% ?

Exercice 2 : calcul de π



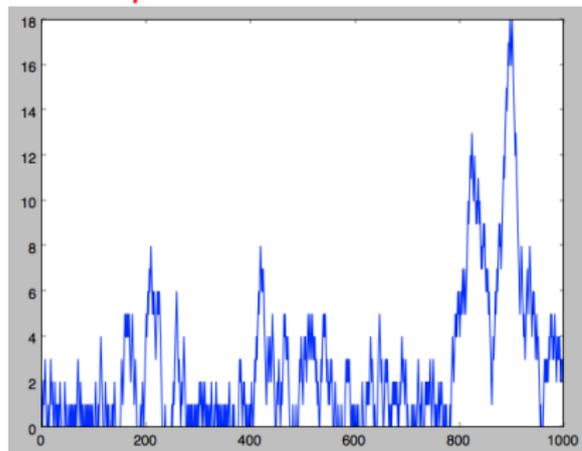
- ▶ La probabilité de tomber sur une arrête est $2/\pi \approx 0.64$.
- ▶ Combien faut-il faire d'expériences pour garantir k chiffres après la virgules ?

Attention aux hypothèses... : Moyenne : $\bar{X} \pm 2 \frac{\bar{\sigma}}{\sqrt{n}}$. ?

1. Indépendance.

Attention aux hypothèses... : Moyenne : $\bar{X} \pm 2 \frac{\bar{\sigma}}{\sqrt{n}}$. ?

1. Indépendance.



Ici : moyenne : $\bar{X} = 2.60$, variance normalisée : $\frac{\bar{\sigma}}{\sqrt{n}} = 0.10$.

Vraie moyenne : 2.

Attention aux hypothèses... : Moyenne : $\bar{X} \pm 2 \frac{\bar{\sigma}}{\sqrt{n}}$. ?

1. Indépendance.

2. Problème de variance

Loi de *Zipf* :

$$\mathbb{P}(X_n = i) \propto \frac{1}{i^\alpha}$$

Soit $S_n = \frac{1}{n} \sum_{k=1}^n X_k$.

▶ $(S_n - n\mu)/\sqrt{n}$

Attention aux hypothèses... : Moyenne : $\bar{X} \pm 2\frac{\bar{\sigma}}{\sqrt{n}}$. ?

1. Indépendance.

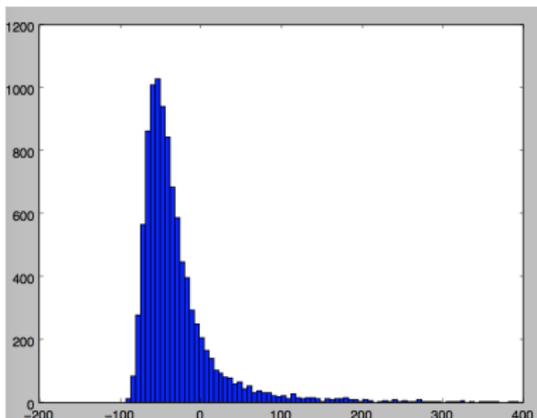
2. Problème de variance

Loi de Zipf :

$$\mathbb{P}(X_n = i) \propto \frac{1}{i^\alpha}$$

Soit $S_n = \frac{1}{n} \sum_{k=1}^n X_k$.

- ▶ $(S_n - n\mu)/\sqrt{nex} : (\alpha = 1.2)$
 $n = 1000$



Attention aux hypothèses... : Moyenne : $\bar{X} \pm 2\frac{\bar{\sigma}}{\sqrt{n}}$. ?

1. Indépendance.

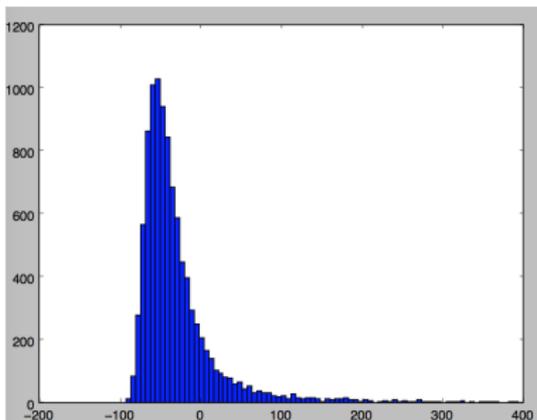
2. Problème de variance

Loi de Zipf :

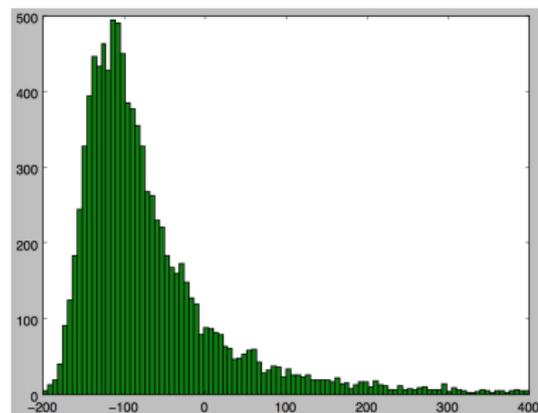
$$\mathbb{P}(X_n = i) \propto \frac{1}{i^\alpha}$$

Soit $S_n = \frac{1}{n} \sum_{k=1}^n X_k$.

► $(S_n - n\mu)/\sqrt{nex} : (\alpha = 1.2)$
 $n = 1000$



$n = 10000$



Les intervalles de confiance sur les quantiles sont plus robustes

(au tableau)

Convergence presque sûre

Définition et exemples

Loi des grands nombres

Modèles de population

Résumé de l'épisode (précédent)

Convergence en loi

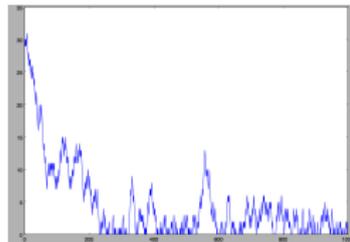
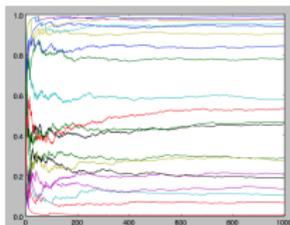
Définition et exemples

Théorème central limite et intervalles de confiance

A retenir

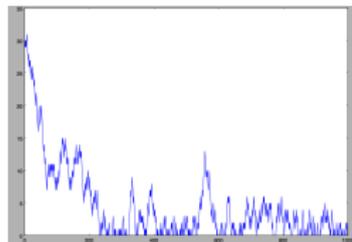
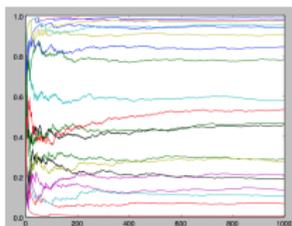
A retenir

Deux types de convergence

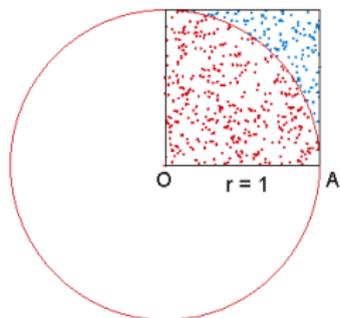


A retenir

Deux types de convergence

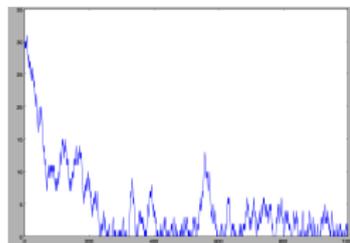
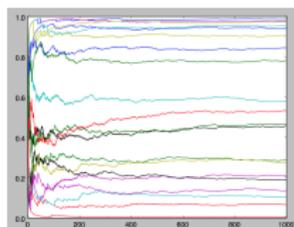


Loi des grands nombres

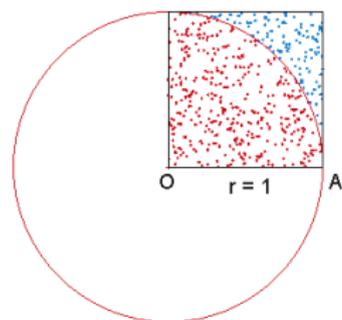


A retenir

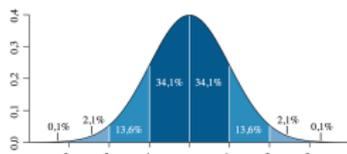
Deux types de convergence



Loi des grands nombres



Théorème centrale limite
intervalles de confiance.



$$\mathbb{E}(\bar{X}) = \bar{X} \pm 2 \frac{\bar{\sigma}}{\sqrt{n}}$$

sous hypothèses