telecom Saint-Etienne

Probabilités et Statistiques

Année 2009/2010

laurent.carraro@telecom-st-etienne.fr

olivier.roustant@emse.fr

Cours n°9

Théorème de la limite centrée

Monte Carlo

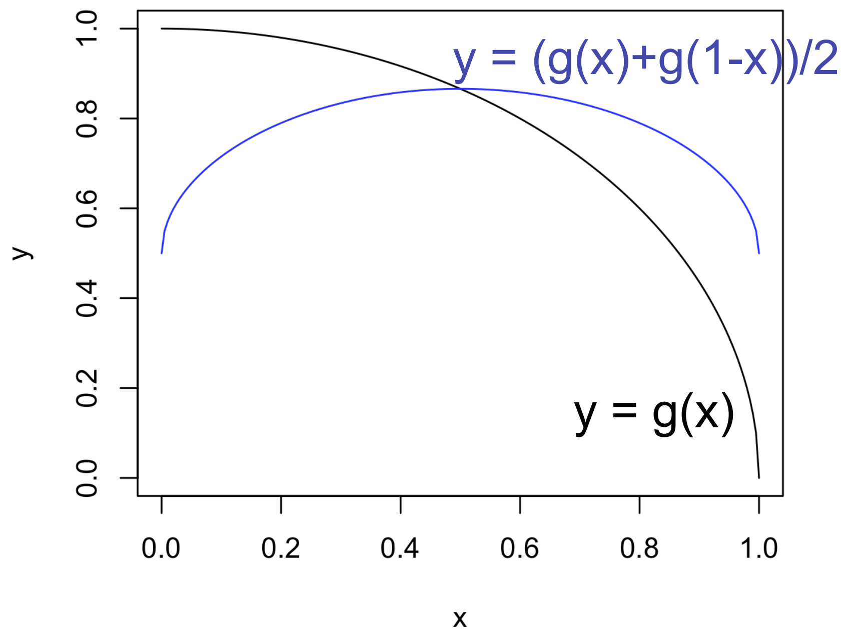
Retour sur l'exercice

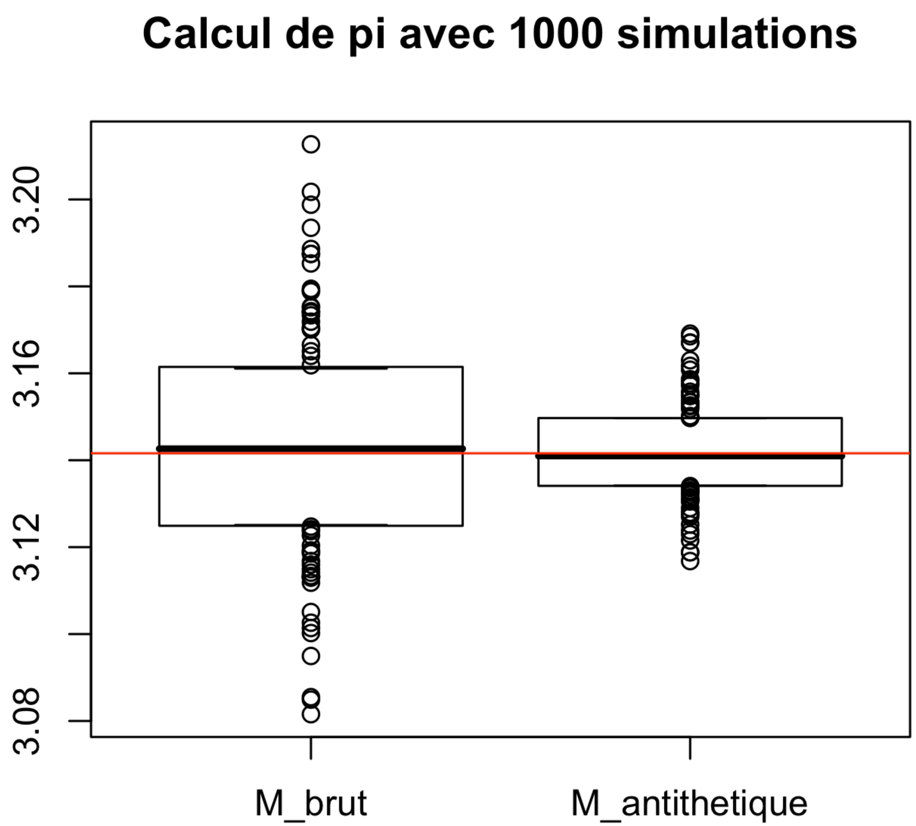
On veut calculer π par simulation, en remarquant que $\pi/4 = E(g(U))$, avec U unif. sur [0,1] et $g(x)=(1-x^2)^{1/2}$ (arc de cercle…)

Estimateur : $M_n = (g(U_1)+\ldots+g(U_n))/n$

Que vaut l'erreur quadratique moyenne dans ce cas ? $E[(M_n - \pi/4)^2] = var(M_n) = var(g(U))/n$

Exercice (illustration)





Boxplots : résultats pour 100 répétitions des 1000 simulations

Questions

On effectue 1000 simulations, on obtient 3.104

- Peut-on donner une idée de l'erreur autour de 3.104 ?
- Combien de simulations faut-il faire en plus pour que le 2ème chiffre soit bon ?

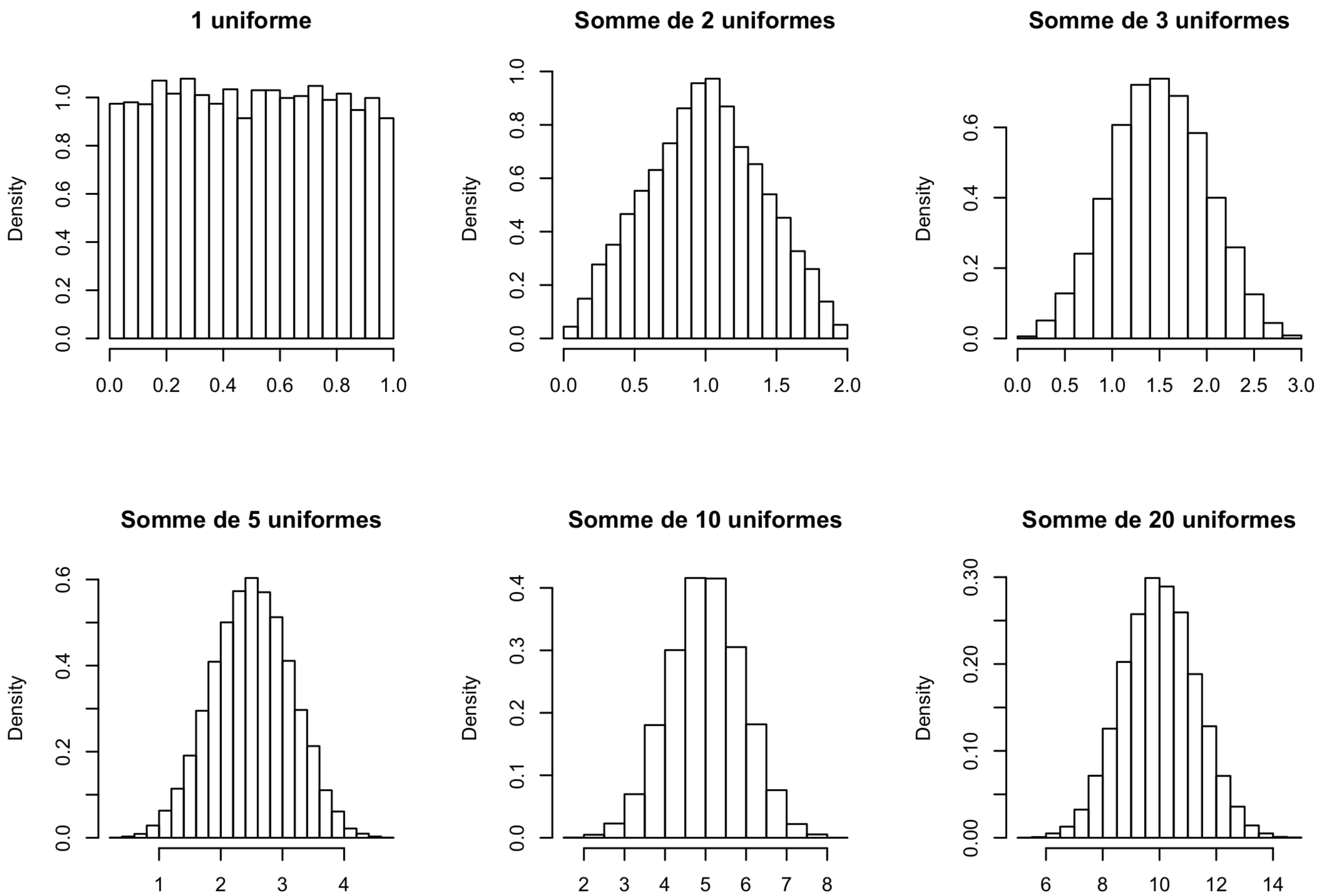
Problématique générale

- THEOREME. [Loi des grands nombres].
  Soient $X_1, \ldots, X_n$ des v.a. i.i.d. telles que $E(X_i)$ existe.
  Soit $\mu := E(X_1)$ et $M_n := (X_1 + \ldots + X_n)/n$.
  Alors $M_n \rightarrow E(X_1)$

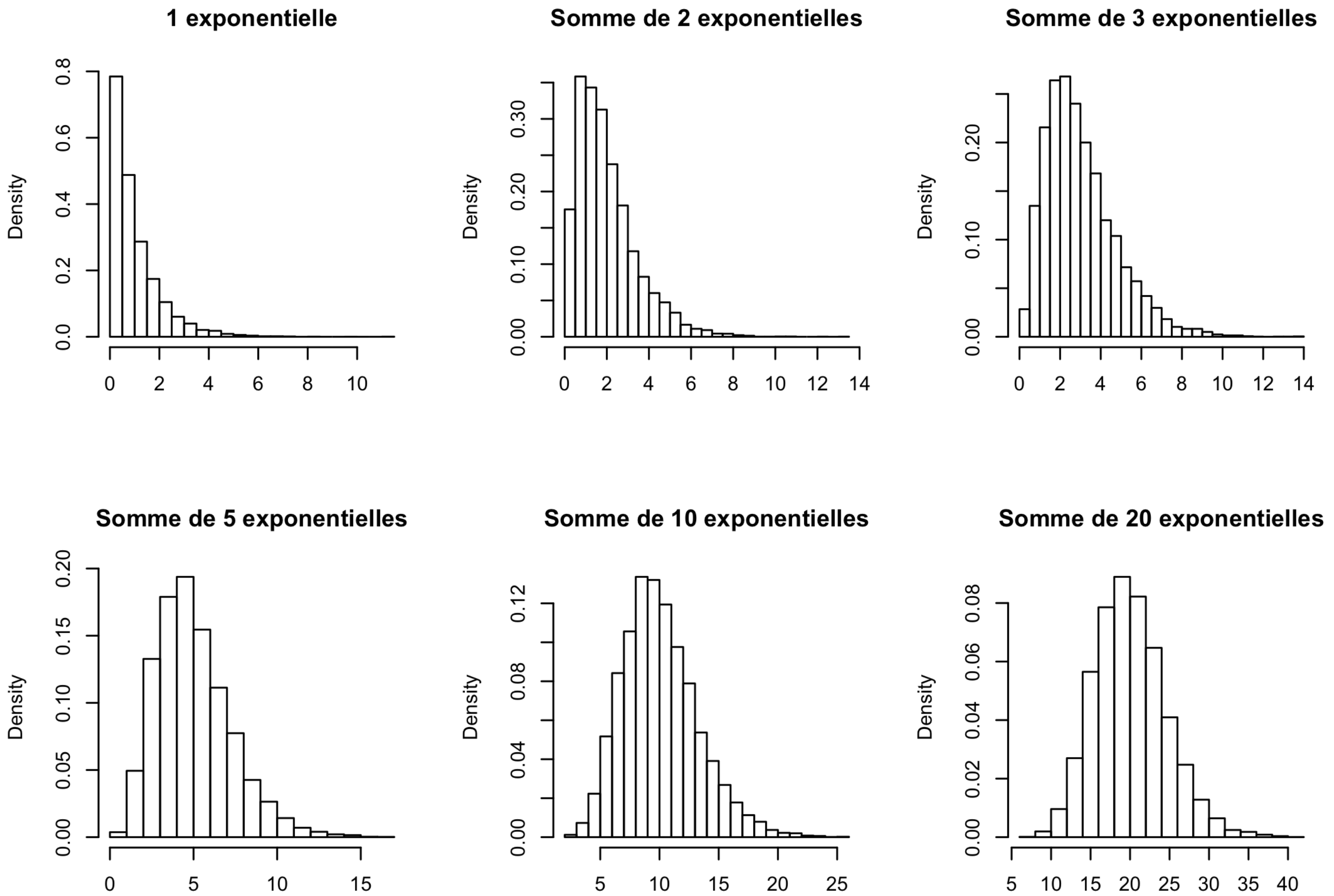
- Erreur quadratique moyenne : $E((M_n - \mu)^2) = var(M_n) = \sigma^2 / n$

- QUESTION.
  Comment se répartit M_n autour de la valeur limite μ ?
  ➔ permettra d'obtenir des intervalles de confiance

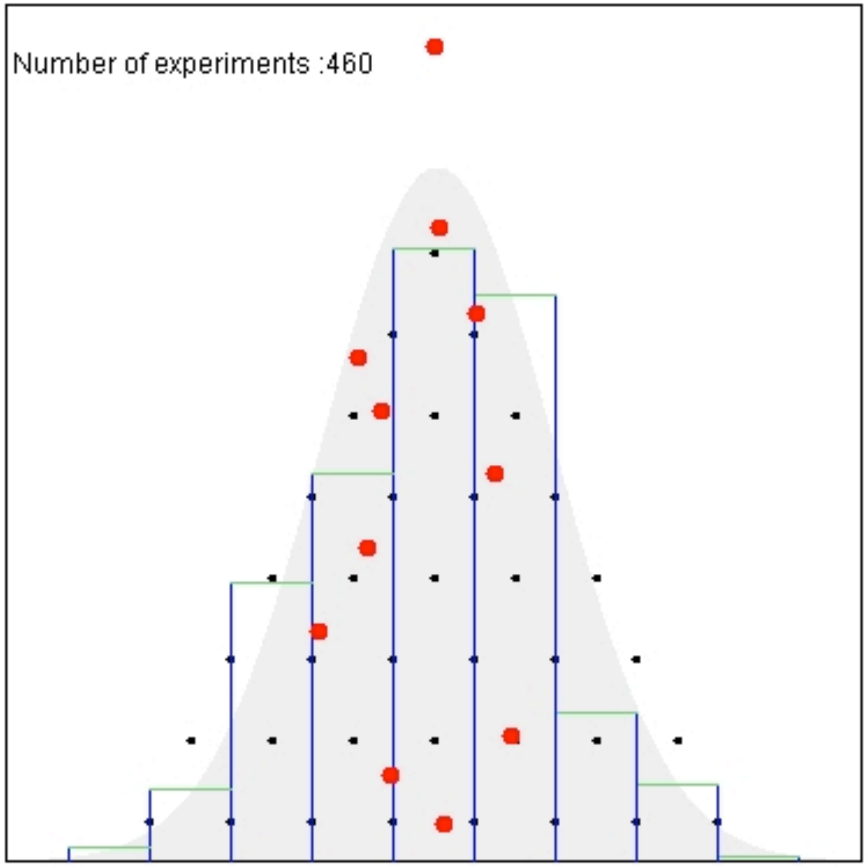
Loi d’une somme de v.a. i.i.d ?



Loi d’une somme de v.a. i.i.d ?



Cas de la loi de Bernoulli planche de Galton



Source : http://www-sop.inria.fr/mefisto/java/tutorial1/node11.html

Planche de Galton : modélisation

- X_i : déplacement de la boule au $i^{ème}$ étage : +1 si elle part à droite, -1 si elle part à gauche. Convention: $X_0 = 0$.
- Hypothèse : les X_i sont 2 à 2 indépendantes.
- Soit $S_n = X_1 + \ldots + X_n$

- *Questions*
  1. Quelle est la loi de X_i ?
  2. Comment s'interprète S_n ?
  3. Traduire en termes probabilistes le phénomène observé sur la planche de Galton

Planche de Galton et loi binomiale

- Définition : la loi binomiale B(n,p) est la loi de $Y_1+\ldots+Y_n$ lorsque $Y_1, \ldots, Y_n$ sont i.i.d. de loi B(p) (de support {0,1}).

- *Questions*
  - Interpréter avec le jeu de pile ou face
  - Calcul de la loi :
    - Que vaut $P(Y_1=\ldots=Y_k=1$ et $Y_{k+1}=\ldots=Y_n=0)$?
    - En déduire $P(Y_1+\ldots+Y_n = k)$, pour $0 \leq k \leq n$.
  - Lien avec la planche de Galton
    - Que dire de la loi $B_{n,\ 1/2}$ lorsque n est grand ?

Théorème de la limite centrée

- THEOREME. [de la limite centrée]
  Soient $X_1, \ldots, X_n$ des v.a. i.i.d. On ne précise pas la loi.
  On suppose seulement que $var(X_i)$ existe.
  Soit $\mu := E(X_1)$, $\sigma^2 := Var(X_1)$ et $M_n := (X_1 + \ldots + X_n)/n$.
  Alors pour n grand, M_n est approx. de loi $N(\mu, \sigma^2/n)$

  Ou encore : $(M_n - \mu)/(\sigma/\sqrt{n})$ est approx. de loi $N(0,1)$

  Formulation rigoureuse [convergence en loi] :
  Pour tout x, $P((M_n - \mu)/(\sigma/\sqrt{n}) < x) \rightarrow F_{N(0,1)}(x)$

Commentaires

- Remarquer la généralité des hypothèses sur la loi des X_i !
  - On peut même alléger l'hypothèse d'indépendance…
- Contre-exemple si var(X_i) n'existe pas.
  - Loi de Cauchy : $f(t) = (1/\pi)\times 1/(1+t^2)$
  - *Proposition : si les X_i sont i.i.d. et de loi de Cauchy, alors pour tout n, M_n est de loi de Cauchy.*

Démonstration

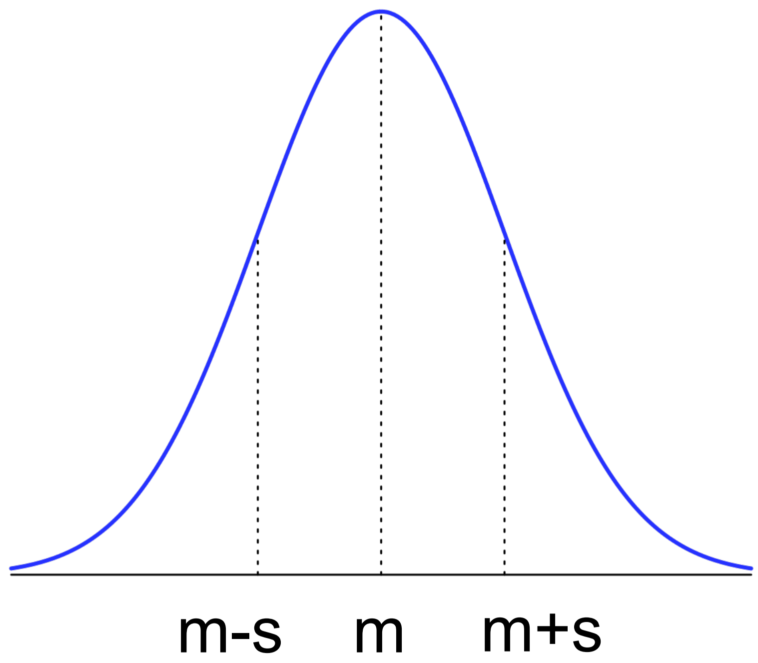
- Historique
  - De Moivre (1733) : cas de la loi B(p) avec p=1/2
  - Laplace (1812) : cas de la loi B(p), p quelconque
  - Lévy (≈1920) : cas général

- Grandes lignes (démo de Lévy, voir polycopié page 80)
  - Utilise la transformée de Fourier Φ d'une v.a. (fonction caractéristique) et ses propriétés.
  - Un calcul simple permet de voir que la fonction caract. de $(M_n - \mu)/(\sigma/\sqrt{n})$ tend vers la fonction $t \rightarrow \exp(-t^2/2)$
  - Or $t \rightarrow \exp(-t^2/2)$ est la fonction caract. de la loi N(0,1)

Loi normale

- La loi N(m, s^2) est définie par sa densité

$$\frac{1}{s\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-m}{s}\right)^2\right)$$



Loi normale standard

- On peut toujours se ramener à la loi $N(0,1)$
  = loi normale standard

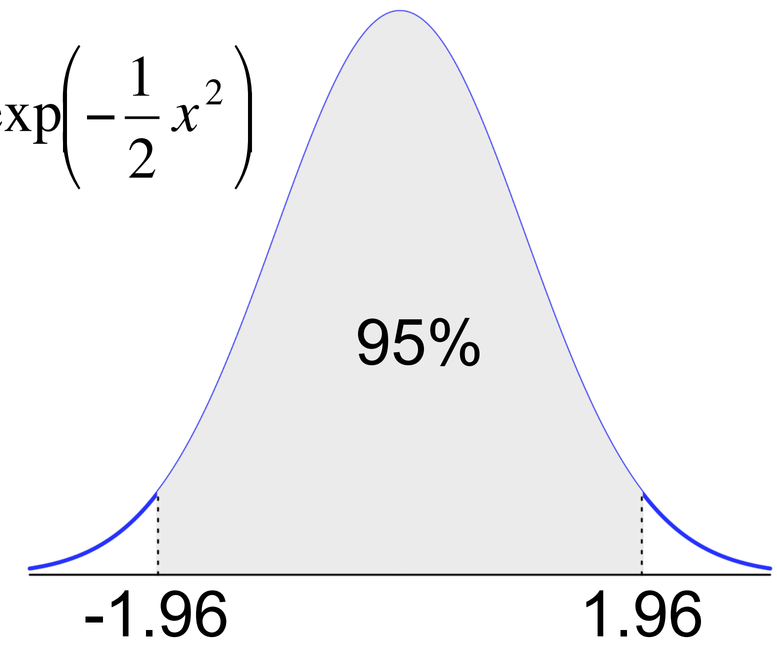
EXERCICE

- Soit X de loi $N(0, 1)$. Quelle est la loi de $m+sX$?
- Calculer $E(X)$ et $var(X)$
- Que vaut E et var pour une loi $N(m, s^2)$?
- Comment passer d'une v.a. de loi $N(m,s^2)$ à une v.a. de loi $N(0,1)$?

Loi normale standard

- Un intervalle à connaître

$$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right)$$



Application : intervalle de confiance

- Déduire du TLC que pour n « grand »:
  $P([M_n - 2\sigma/\sqrt{n} \,;\, M_n + 2\sigma/\sqrt{n}]$ contient $\mu) \approx 95\%$

- On admet que le résultat reste vrai en remplaçant σ^2 par un de ses 2 estimateurs (cf. polycopié p. 65):
  $\sigma_X^2 := 1/n \times \Sigma(X_i - M_n)^2$ ou $S_X^2 := 1/(n-1) \times \Sigma(X_i - M_n)^2$

- Vocabulaire : pour n assez grand,
  $[M_n - 2S_X/\sqrt{n} \,;\, M_n + 2S_X/\sqrt{n}]$ est un int. de conf. à 95% de μ

Retour sur l'exemple

- Exemple avec n = 1000 simulations

Moyenne empirique (M_n) : 3.104
Ecart-type empirique (σ_X) : 0.881

Intervalle de confiance : [3.048 3.159]