telecom
Saint-Etienne

Probabilités et Statistiques

Année 2010/2011

laurent.carraro@telecom-st-etienne.fr

olivier.roustant@emse.fr

Cours n°9

Loi normale

Théorème de la limite centrée

Application – intervalle de confiance

telecom Saint-Etienne

Préambule

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

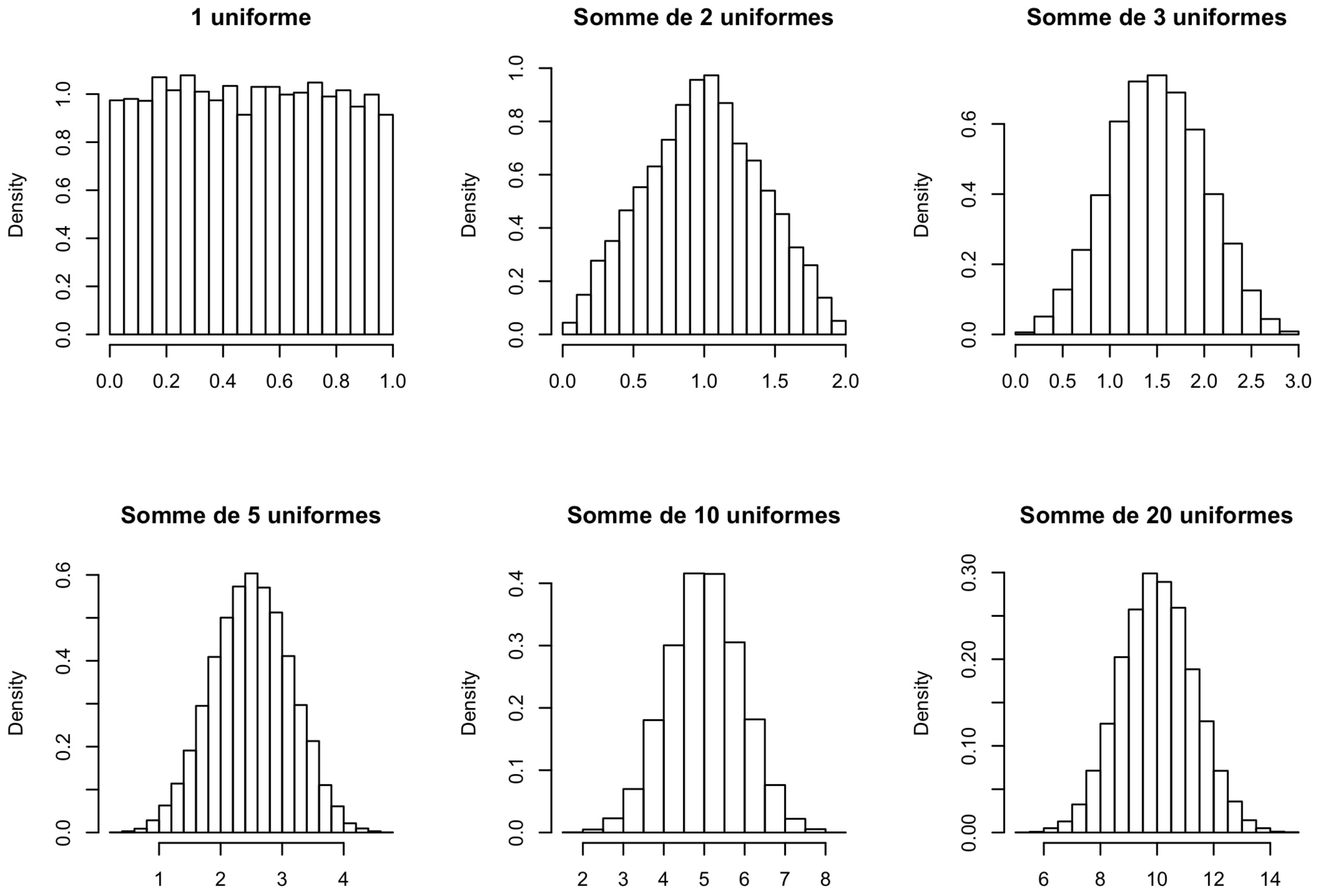
- Exemple. Estimation d'une proportion
  - Que retrouve t-on lorsque $X_1, \dots, X_n$ sont i.i.d $B(p)$?

Préambule

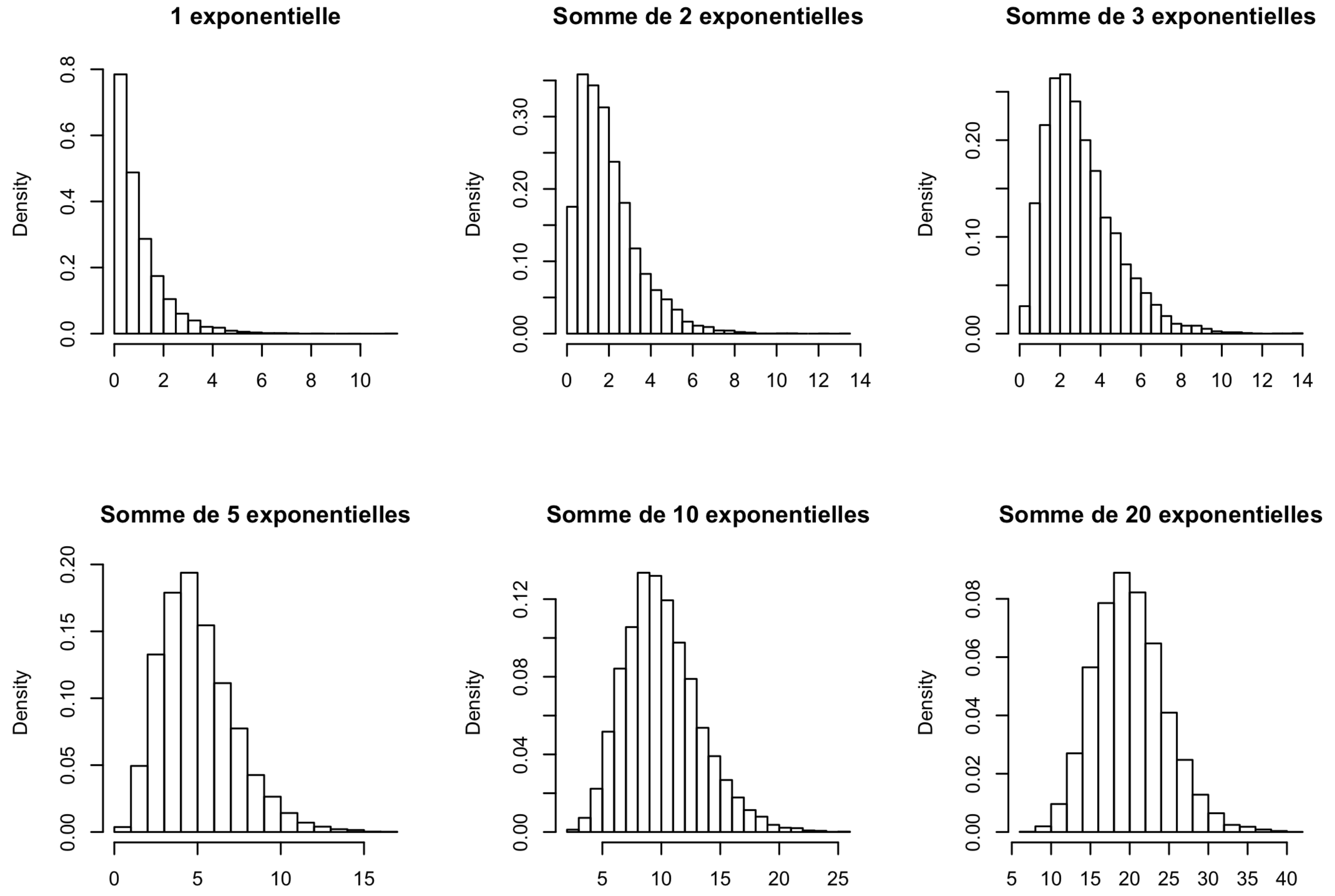
- Vitesse de convergence ?
  Erreur quadratique moyenne :
  $$E((M_n - \mu)^2) = var(M_n) = \sigma^2 / n$$
  -> vitesse en $\boxed{1/\sqrt{n}}$

- 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 ?



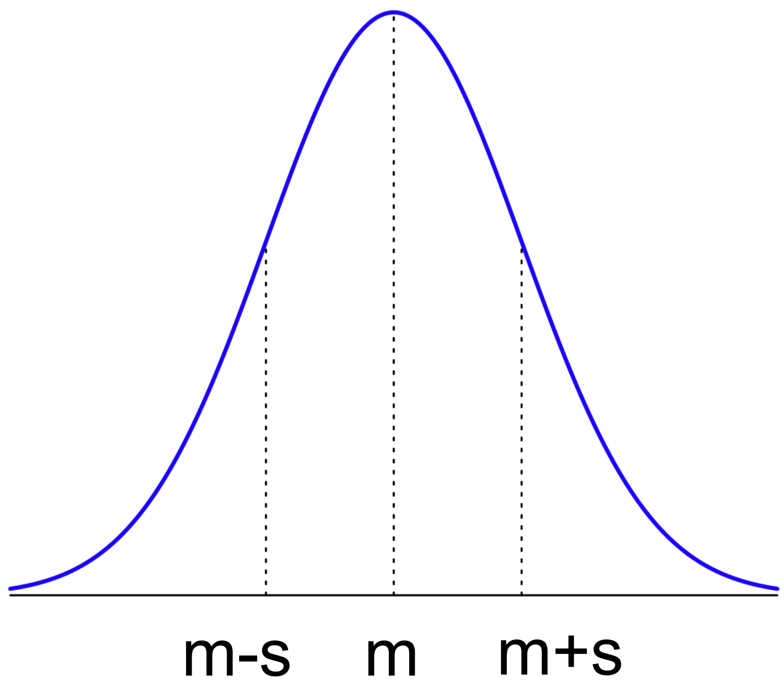
Loi d'une somme de v.a. i.i.d ?



Loi normale

- La loi N(m, s²) 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

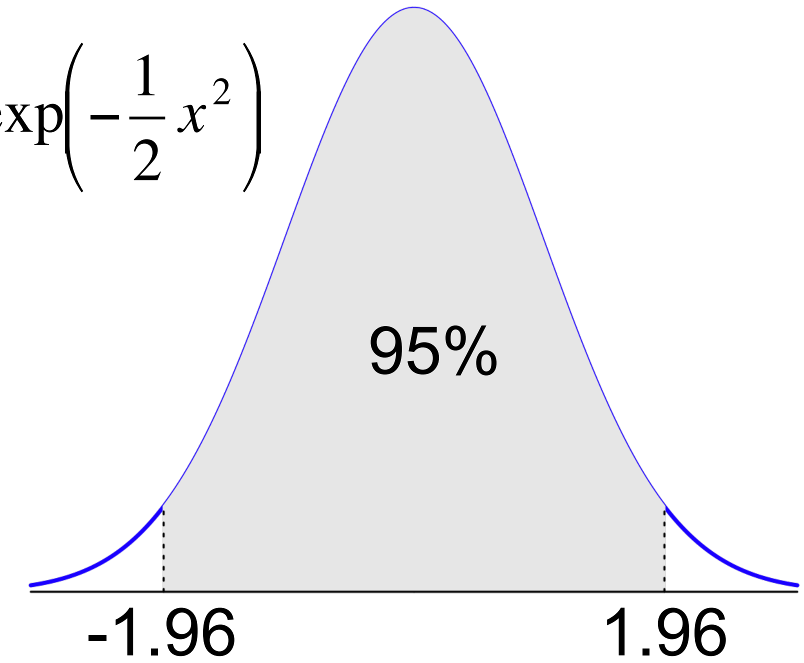
EXERCICE

- Vérifier que $\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$ est une densité de probabilité
- Soit X de loi N(0, 1). Mq m+sX est de loi N(m, s^2)
- Soit Y de loi N(m, s^2). Mq (Y-m)/s est de loi N(0,1)
- Montrer que E(X)=0 et var(X)=1
- En déduire que E(Y)=m, Var(Y)=s^2 si Y est de loi N(m, s^2)

Loi normale standard

- Un intervalle à connaître

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



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 $\mathrm{var}(X_i)$ existe.
  Soit $\mu := E(X_1)$, $\sigma^2 := \mathrm{Var}(X_1)$ et $M_n := (X_1 + \ldots + X_n)/n$.

  Alors pour n grand, M_n est approx. de loi $N(\mu, \sigma^2/n)$

telecom Saint-Etienne

Commentaires

- Remarquer la généralité des hypothèses sur la loi des X_i !
  - On peut même alléger l'hypothèse d'indépendance…

- *Formulation rigoureuse [convergence en loi] :*

  Pour tout x, $P\left((M_n - \mu)/(\sigma/\sqrt{n}) < x \right) \rightarrow F_{N(0,1)}(x)$

- 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 (à suivre)

- 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)

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 μ) ≈ 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

Application : intervalle de confiance

- ILLUSTRATION.

  Vous êtes 2. A tour de rôle, vous lancez 100 fois la même pièce de monnaie (équilibrée ?).

  On compte 1 si pile, 0 si face. On obtient :

  - Moyenne des 100 simulations : 0.54 (joueur 1), 0.43 (joueur 2)
  - Ecart-type des 100 simulations : 0.50 (joueur 1), 0.50 (joueur 2)

  Soit p la probabilité d'obtenir "pile".

  Intervalle de confiance approché à 95% de p ?

Application : intervalle de confiance

- ILLUSTRATION (suite)

  Vous êtes 100 joueurs. Que pouvez-vous dire des 100 intervalles de confiance que vous obtenez ?

  Est-ce qu'ils contiennent p ?

Application : intervalle de confiance

INTERPRETATION

- La proportion des intervalles de confiance à 95% qui contiennent μ tend vers 95%