Analyse Numérique - 5

Jacques Le Bourlot
Observatoire de Paris & Université Paris-Diderot

13 Janvier 2014

Moments d'une distribution

Étant donnés une fonction de distribution $f(x)$, on a :

- $Prob\,(x_0 < x < x_0 + dx) = f(0)\,dx$
- $\int_{-\infty}^{+\infty} f(x)\,dx = 1$

On appelle Moment d'ordre n de f la quantité :

$$m_n = \int_{-\infty}^{\infty} x^n\,f(x)\,dx$$

Noter que m_n n'existe pas forcément pour tout n.

Évaluation à partir d'observations

On dispose de N valeurs mesurées de x, comment évaluer $f(x)$ d'une part et les m_n de l'autre ?

- Moyenne : Estimateur de m_1 :

$$\tilde{m}_1 = \frac{1}{N} \sum_i x_i$$

Évaluation à partir d'observations

On dispose de N valeurs mesurées de x, comment évaluer $f(x)$ d'une part et les m_n de l'autre ?

- Moyenne : Estimateur de m_1 :

$$\tilde{m}_1 = \frac{1}{N} \sum_i x_i$$

- Variance/Écart-type : Estimateur de m_2 :

$$\tilde{m}_2 = \frac{1}{N-1} \sum_i (x_i - m_1)^2 \;; \qquad \sigma = \sqrt{m_2}$$

Évaluation à partir d'observations

On dispose de N valeurs mesurées de x, comment évaluer $f(x)$ d'une part et les m_n de l'autre ?

- Moyenne : Estimateur de m_1 :

$$\tilde{m}_1 = \frac{1}{N} \sum_i x_i$$

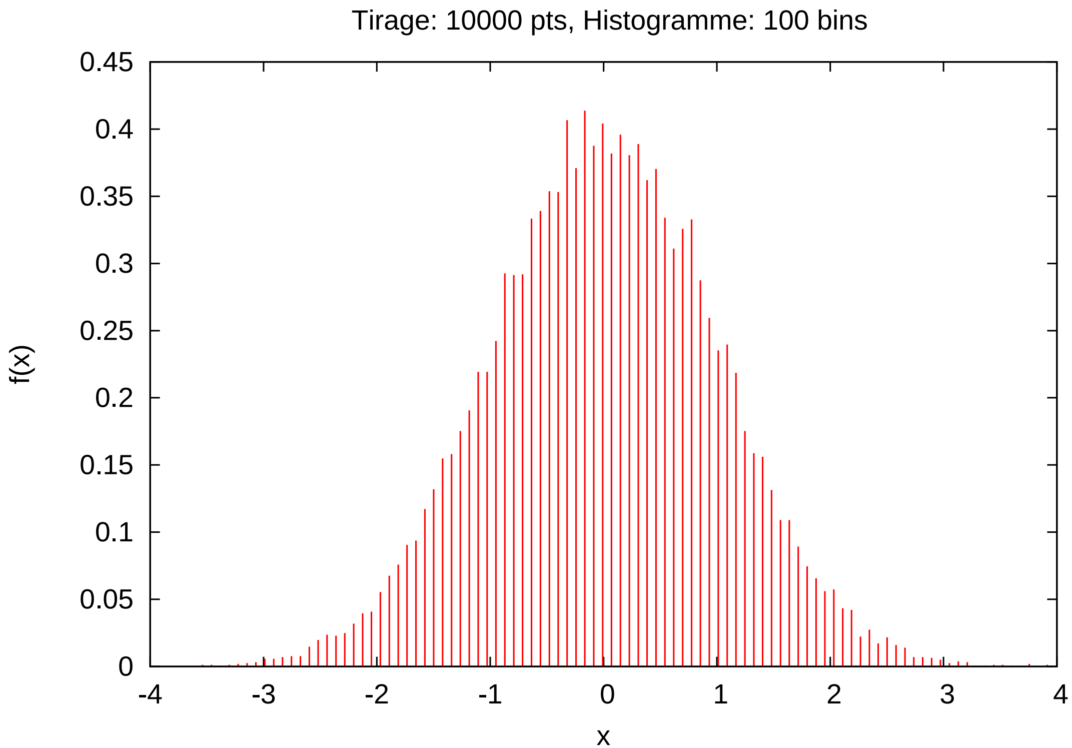
- Variance/Écart-type : Estimateur de m_2 :

$$\tilde{m}_2 = \frac{1}{N-1} \sum_i (x_i - m_1)^2 \; ; \qquad \sigma = \sqrt{m_2}$$

- Pour les moments d'ordre supérieur, l'incertitude croit vite (on donne du poids aux points peu échantillonnés).

Fonction de distribution

- Pour $f(x)$, il faut “biner” :
  - Faible erreur : il faut des pas larges
  - Bonne résolution il faut des pas petits



Fonction de répartition cumulée

- Solution : Il faut utiliser la fonction de répartition cumulée ! :
  - c'est la "bonne" grandeur statistique
  - Définie par :

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$

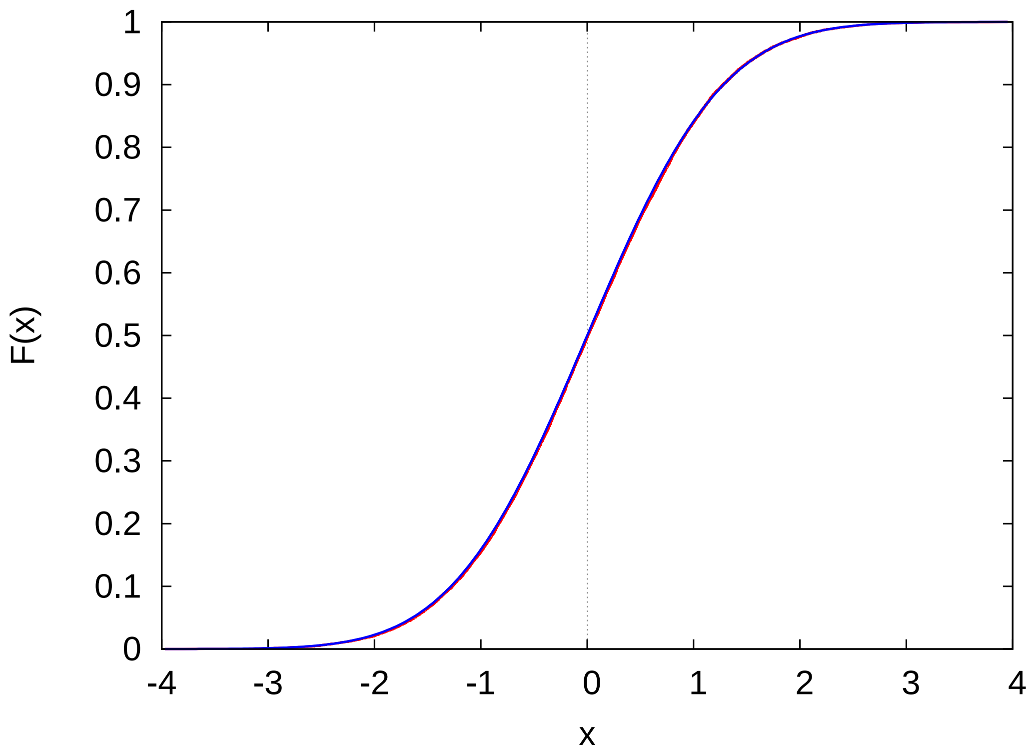
  - Propriété :

$$Pr(x < x_0) = F(x_0)$$

  - On l'obtient par un tri des valeurs des x_i observés.

$$Pr(x < x_i) \simeq \frac{i}{N}$$

Fonction de répartition cumulée



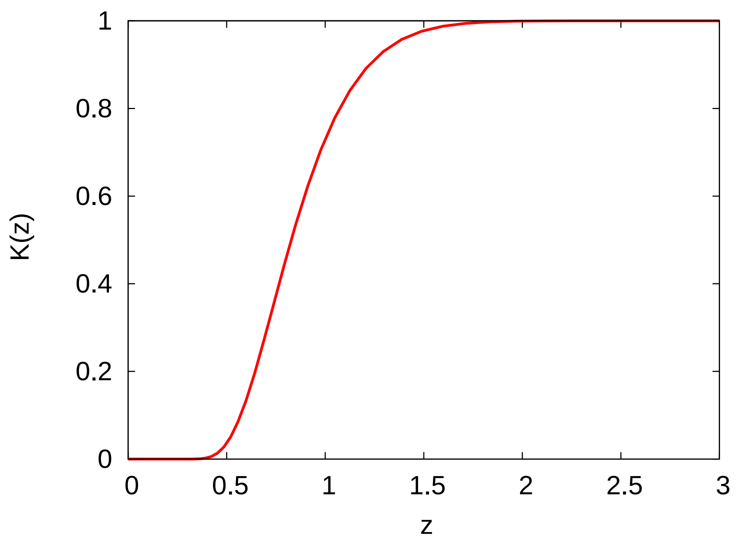
Gaussienne : $F(x) = \frac{1}{2}\left[\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)+1\right]$

Tests Statistiques

Tirage suivant distribution donnée - Kolmogorov :

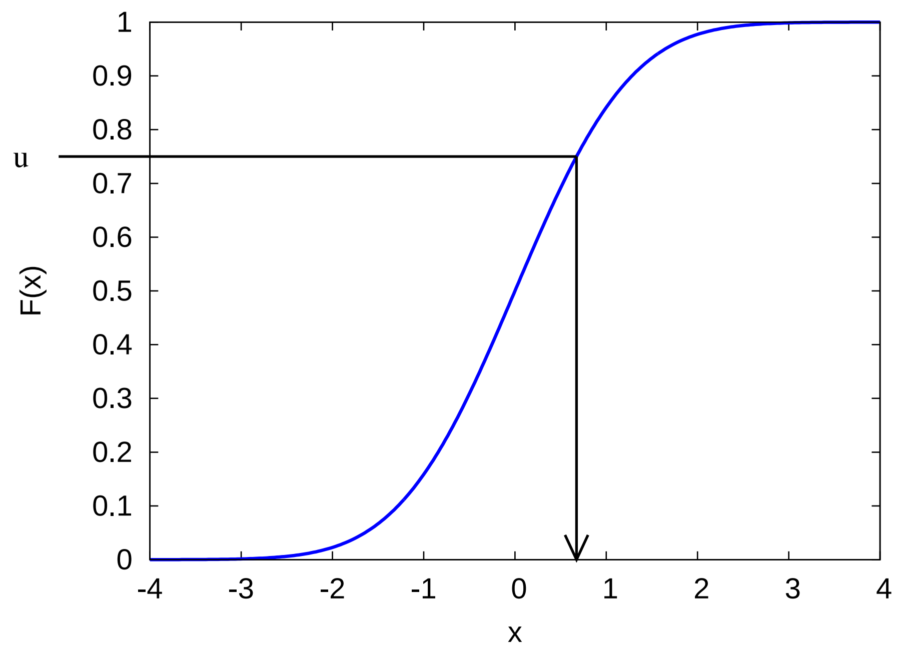
$$Pr\left(\sqrt{n}\,\sup_n |F_n(x) - F(x)| > \alpha\right) = K(\alpha)$$

$$K(x) = 2\sum_{j=1}^{\infty}(-1)^{j-1}\, e^{-2j^2x^2} = \frac{\sqrt{2\pi}}{x}\sum_{j=1}^{\infty}\exp\left(-\frac{(2j-1)^2\,\pi^2}{8x^2}\right)$$



Génération de nombres aléatoires

Pour tirer un nombre suivant la loi $f(x)$, on part de la fonction de répartition $F(x)$.



$$\frac{du}{dx} = f(x)$$

Génération de nombres aléatoires

Si U est tiré suivant une loi uniforme (entre 0 et 1), alors $x = F^{-1}(U)$ est distribué suivant $f(x)$.

- Preuve :

$$
\begin{aligned}
& Pr(F^{-1}(U) < x) \\
&= Pr\left(\{y | F(y) = U\} < x\right) \\
&= Pr\left(U < F(x)\right) \\
&= F(x)
\end{aligned}
$$

- Utilise :
  - Le fait que F est strictement croissante
  - Le fait que U est uniforme sur $[0 : 1]$

Ajustement

Étant donnés des points expérimentaux (x_i, y_i, σ_i) où σ_i est l'incertitude sur y_i, trouver le meilleur jeux de paramètres $\{a_j\}$ tels que :

$$y = f_{\{a_j\}}(x)$$

On forme :

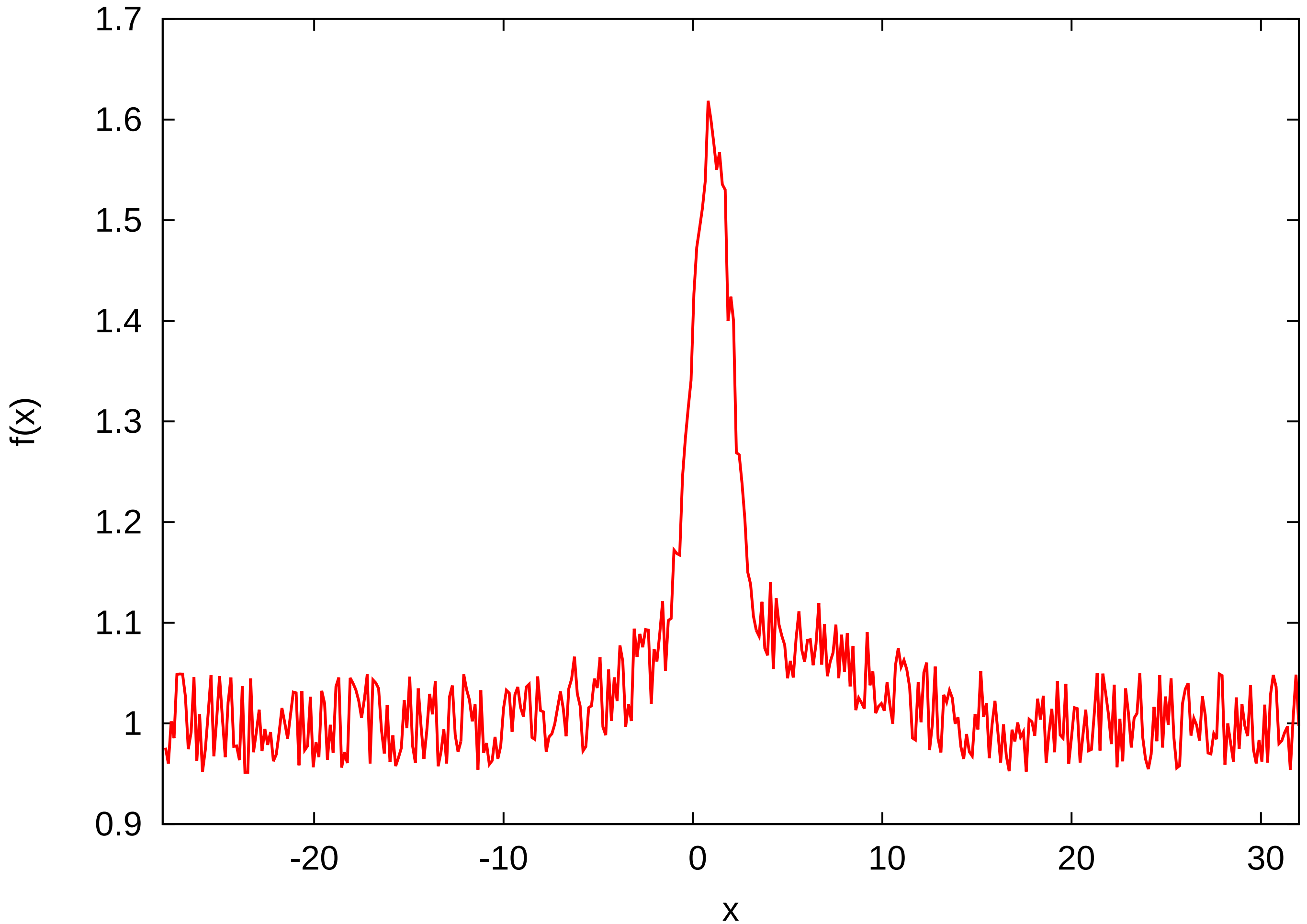
$$\chi^2(a_j) = \sum_i \left(\frac{f(x_i) - y_i}{\sigma^2} \right)^2$$

et on cherche à minimiser $\chi^2(a_j)$:

$$\frac{\partial \chi^2}{\partial a_j} = 0$$

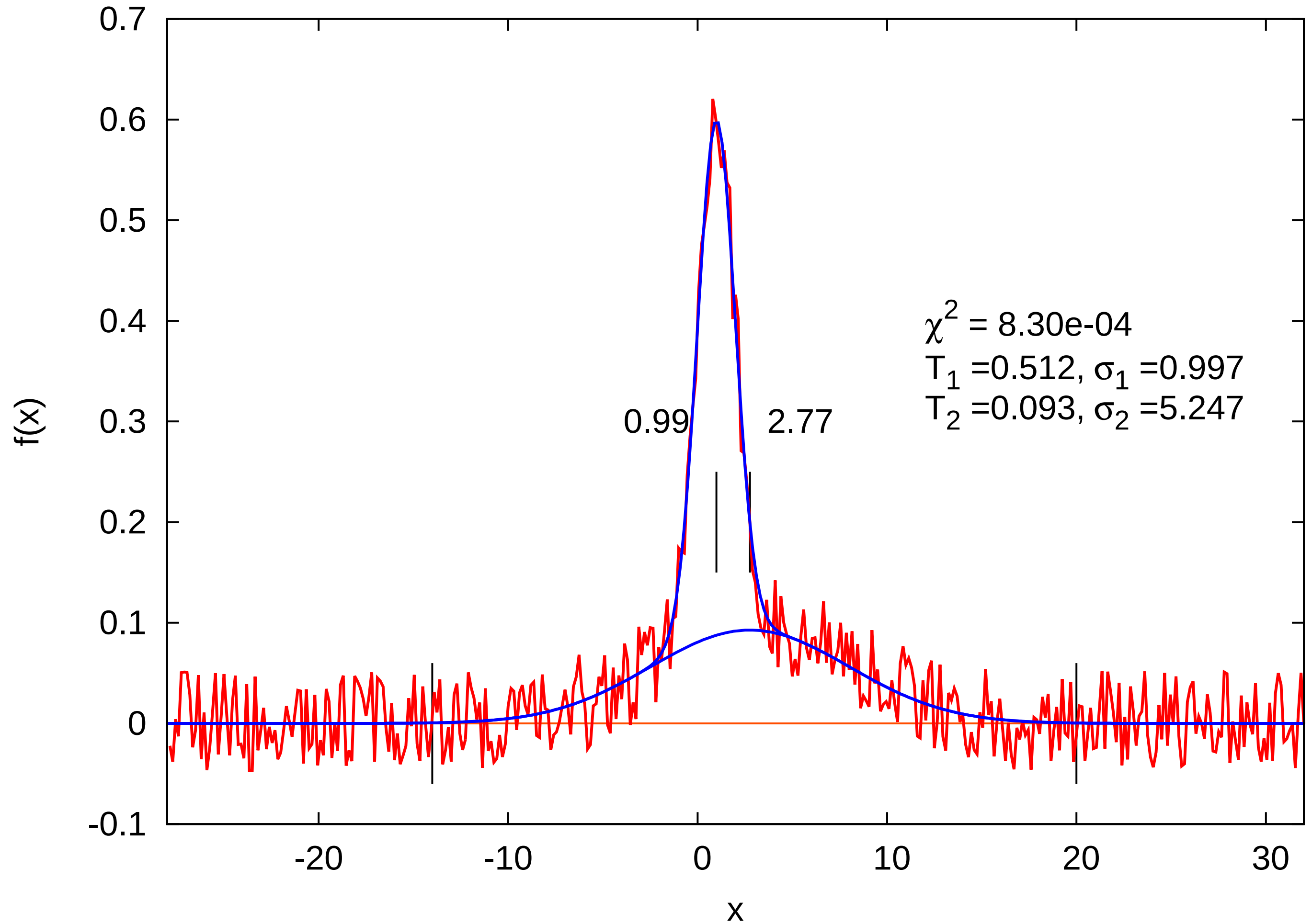
Si f est linéaire par rapport aux $a_j \Rightarrow$ ajustement de moindres carrés classique (solution analytique)
Sinon : Algorithme de Levenberg-Marquardt

Ajustement - Exemple



Voir l'exemple "gnuplot" complet.

Ajustement - Exemple



Voir l'exemple "gnuplot" complet.

Transformée de Fourier rapide

- Algorithme papillon : Superbe, mais impossible à programmer $\Rightarrow$ Utiliser une bibliothèque !
- Relation résolution <-> Valeurs maximales

$$\Delta\nu = \frac{1}{T_{max}}\,; \quad \nu_{max} = \frac{1}{2\,\Delta t}$$

- Théorème de Parseval :

$$2\pi \int_{-\infty}^{\infty} \left|\tilde{f}(\nu)\right|^2 \, d\nu = \int_{-\infty}^{\infty} |f(t)|^2 \, dt$$

- Shanon : Pour conserver toute l'information, il faut échantillonner à $2\,\nu_{max}$.

Diagonalisation

- Calcul de la plus grande valeur propre.
- Calcul de toutes les valeurs propres.
- Calcul des valeurs propres ET vecteurs propres.
- Cas des matrices symétriques.