Statistique

Corinne Mailhes et Jean-Yves Tourneret[1]

(1) Université de Toulouse, ENSEEIHT-IRIT-TeSA

Corinne.Mailhes@tesa.prd.fr et jyt@n7.fr

Plan du cours

- Chapitre 1 : Convergence et théorèmes limites
- Chapitre 2 : Estimation
- Chapitre 3 : Tests Statistiques

Bibliographie

- B. Lacaze, M. Maubourguet, C. Mailhes et J.-Y. Tourneret, Probabilités et Statistique appliquées, Cépadues, 1997.
- Athanasios Papoulis and S. Unnikrishna Pillai, Probability, Random Variable and Stochastic Processes, McGraw Hill Higher Education, 4th edition, 2002.

Convergence en loi

- **Définition**
  La suite de va $X_1, ..., X_n$ converge en loi vers la va X si et ssi la suite des fonctions de répartition $F_n(x) = P[X_n < x]$ converge simplement vers $F(x) = P[X < x]$ en tout point x où F est continue.

- **Notation**

$$X_n \underset{n \to \infty}{\overset{\mathcal{L}}{\to}} X$$

- **Exemple**

$$P[X_n = 1] = \frac{1}{n} \text{ et } P[X_n = 0] = 1 - \frac{1}{n}$$

Convergence en loi

- Propriétés
  - Théorème de Levy
    X_n cv en loi vers X si et ssi ϕ continue en $t = 0$ et

$$\phi_n(t) = E\left[e^{itX_n}\right] \underset{n \to \infty}{\rightarrow} \phi(t) = E\left[e^{itX}\right], \forall t.$$

  - Si X_n est une suite de va continues de densités $p_n(x)$ et que $p_n(x) \underset{n \to \infty}{\rightarrow} p(x)$ p.p., alors $X_n \underset{n \to \infty}{\overset{\mathcal{L}}{\rightarrow}} X$.
  - Si $X_n \underset{n \to \infty}{\overset{\mathcal{L}}{\rightarrow}} X$ et $g : \mathbb{R} \rightarrow \mathbb{R}$ continue, alors

$$g(X_n) \underset{n \to \infty}{\overset{\mathcal{L}}{\rightarrow}} g(X).$$

Convergence en probabilité

- Définition
  La suite de va $X_1, ..., X_n$ converge en probabilité vers la va X si et ssi $\forall \epsilon > 0$, on a

$$P\left[|X_n - X| > \epsilon\right] \underset{n\to\infty}{\rightarrow} 0.$$

- Notation

$$X_n \underset{n\to\infty}{\overset{\mathcal{P}}{\rightarrow}} X$$

- Exemple : X_n de densité $p_n(x) = \frac{ne^{-nx}}{(1+e^{-nx})^2}$.
- Propriété
  Si $X_n \underset{n\to\infty}{\overset{\mathcal{P}}{\rightarrow}} X$ et $g : \mathbb{R} \rightarrow \mathbb{R}$ continue, alors

$$g(X_n) \underset{n\to\infty}{\overset{\mathcal{P}}{\rightarrow}} g(X).$$

Convergence en moyenne quadratique

- Définition

  La suite de va $X_1, ..., X_n$ converge en moyenne quadratique vers la va X si et ssi

$$E\left[(X_n - X)^2\right] \underset{n\to\infty}{\to} 0.$$

- Notation

$$X_n \underset{n\to\infty}{\overset{\mathcal{MQ}}{\to}} X$$

- Exemple

$$P[X_n = n] = \frac{1}{n^p} \text{ et } P[X_n = 0] = 1 - \frac{1}{n^p}$$

avec $p = 2$ et $p = 3$.

Convergence presque sûre

- Définition
  La suite de va $X_1, ..., X_n$ converge presque sûrement vers la va X si et ssi

$$X_n(\omega) \underset{n\to\infty}{\to} X(\omega), \quad \forall \omega \in A | P(A) = 1.$$

- Notation

$$X_n \underset{n\to\infty}{\overset{\mathcal{PS}}{\to}} X$$

- Comparaison entre les différents types de convergence

Loi faible des grands nombres

- Loi faible des grands nombres
  Si $X_1, ..., X_n$ sont des va indépendantes et de même loi de moyenne $E[X_k] = m < \infty$, alors la va $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ converge en probabilité vers m.
- Preuve

$$\varphi_{\overline{X}_n}(t) = E\left[e^{it\frac{1}{n}\sum_{k=1}^{n} X_k}\right] = E\left[\prod_{k=1}^{n} e^{i\frac{t}{n}X_k}\right] = \left[\varphi\left(\frac{t}{n}\right)\right]^n$$

Dév. de Taylor de ϕ

$$\varphi(t) = \varphi(0) + t\varphi'(0) + t\lambda(t) = 1 + itm + t\lambda(t)$$

Preuve

On en déduit

$$\ln\left[\varphi_{\overline{X}_n}(t)\right] = n \ln\left[1 + i\frac{t}{n}m + \frac{t}{n}\lambda\left(\frac{t}{n}\right)\right] = n\left[i\frac{t}{n}m + \frac{t}{n}\lambda\left(\frac{t}{n}\right)\right]$$

d'où

$$\lim_{n\to\infty} \varphi_{\overline{X}_n}(t) = e^{itm} \qquad \forall t$$

i.e.,

$$\overline{X}_n \underset{n\to\infty}{\overset{\mathcal{L}}{\to}} m \Leftrightarrow \overline{X}_n \underset{n\to\infty}{\overset{\mathcal{P}}{\to}} m$$

Loi forte des grands nombres

- Loi forte des grands nombres
  Si $X_1, ..., X_n$ sont des va indépendantes et de même loi de moyenne $E[X_k] = m < \infty$ et de variance $\sigma^2 < \infty$, alors la va $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ converge en moyenne quadratique vers m.
- Preuve

$$E\left[\left(\overline{X} - m\right)^2\right] = \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n} E\left[\left(X_k - m\right)\left(X_l - m\right)\right]$$

Mais

$$E\left[\left(X_k - m\right)\left(X_l - m\right)\right] = \begin{cases} \sigma^2 \text{ si } k = l \\ 0 \text{ si } k \neq l \end{cases}$$

Preuve

Donc

$$E\left[\left(\overline{X}_n - m\right)^2\right] = \frac{\sigma^2}{n} \underset{n\to\infty}{\to} 0$$

i.e.,

$$\overline{X}_n \underset{n\to\infty}{\overset{\mathcal{MQ}}{\to}} m$$

Théorème de la limite centrale

- Théorème de la limite centrale
  Si $X_1, ..., X_n$ sont des va indépendantes et de même loi de moyenne $E\left[X_k\right] = m < \infty$ et de variance $\sigma^2 < \infty$, alors la va centrée réduite $Y_n = \frac{\sum_{k=1}^{n} X_k - nm}{\sqrt{n\sigma^2}}$ converge en loi vers la loi normale $\mathcal{N}(0,1)$.
- Preuve

$$\varphi_{Y_n}(t) = E\left[e^{itY_n}\right] = e^{-\frac{itm\sqrt{n}}{\sigma}} \prod_{k=1}^{n} E\left[e^{i\frac{t}{\sigma\sqrt{n}}X_k}\right]$$

Mais

$$E\left[e^{i\frac{t}{\sigma\sqrt{n}}X_k}\right] = \varphi\left(\frac{t}{\sigma\sqrt{n}}\right)$$

Preuve

Donc

$$\ln\left[\varphi_{Y_n}(t)\right] = -\frac{itm\sqrt{n}}{\sigma} + n\ln\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)$$

En utilisant le développement de Taylor de φ

$$\varphi(t) = \varphi(0) + t\varphi'(0) + \frac{t^2}{2}\varphi''(0) + t^2\lambda(t)$$

On en déduit

$$\ln\left[\varphi_{Y_n}(t)\right] = -\frac{t^2}{2} + \frac{t^2}{n}\lambda\left(\frac{t}{\sigma\sqrt{n}}\right)$$

$$\lim_{n\to\infty}\varphi_{Y_n}(t) = e^{-\frac{t^2}{2}} \qquad \forall t \;\Leftrightarrow\; Y_n \underset{n\to\infty}{\overset{\mathcal{L}}{\to}} \mathcal{N}(0,1)$$

Que faut-il savoir ?

- Convergence en loi ?
- Convergence en moyenne quadratique ?
- $\frac{1}{n}\sum_{k=1}^{n} X_k$ converge en probabilité vers ? Conditions ?
- $\frac{1}{n}\sum_{k=1}^{n} X_k$ converge en moyenne quadratique vers ? Conditions ?
- $Y_n = \frac{\sum_{k=1}^{n} X_k - ?}{?}$ converge en loi vers ?

Plan du cours

- Chapitre 1 : Convergence et théorèmes limites
- Chapitre 2 : Estimation
  - Modèle statistique, qualités d'un estimateur, exemples
  - Inégalité de Cramér Rao
  - Maximum de vraisemblance
  - Méthode des moments
  - Estimation Bayésienne
  - Intervalles de confiance
- Chapitre 3 : Tests Statistiques

Modèle Statistique

- Observations

$$x_1, ..., x_n$$

- Échantillon

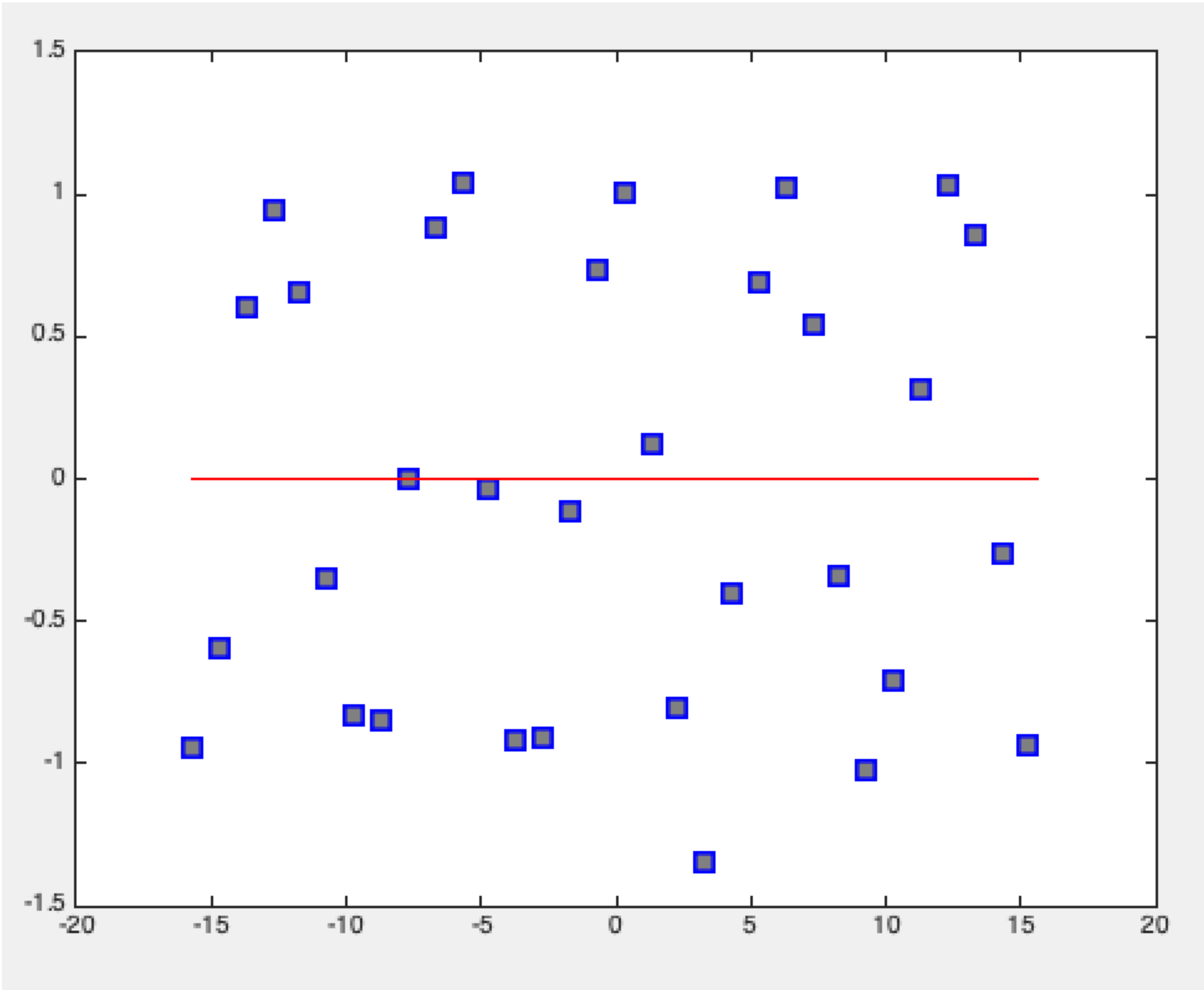
$$X_1, ..., X_n$$

n va iid associées aux observations

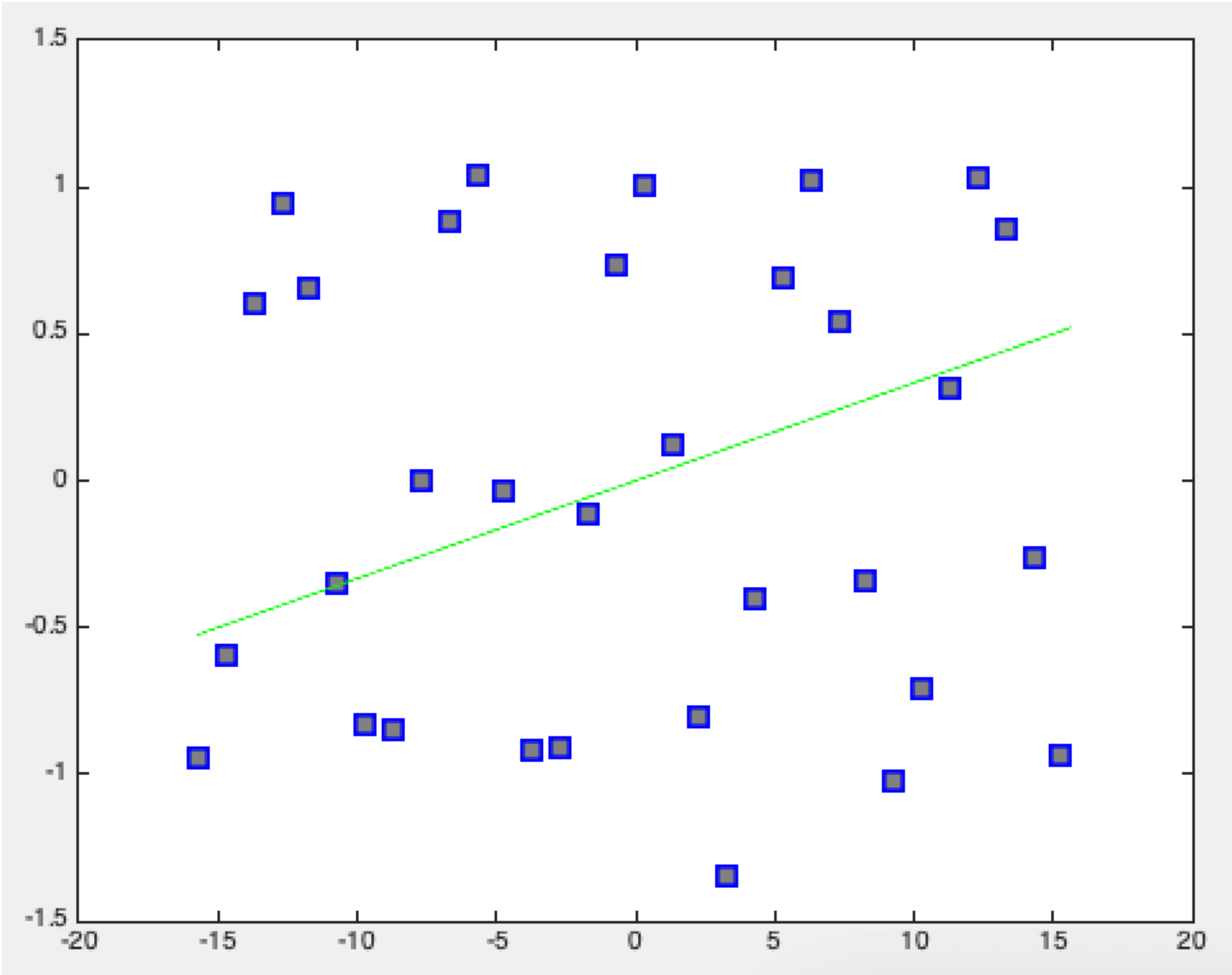
- Estimateur

$$\widehat{\theta}(X_1, ..., X_n) \text{ ou } \widehat{\theta}_n \text{ ou } \widehat{\theta}$$

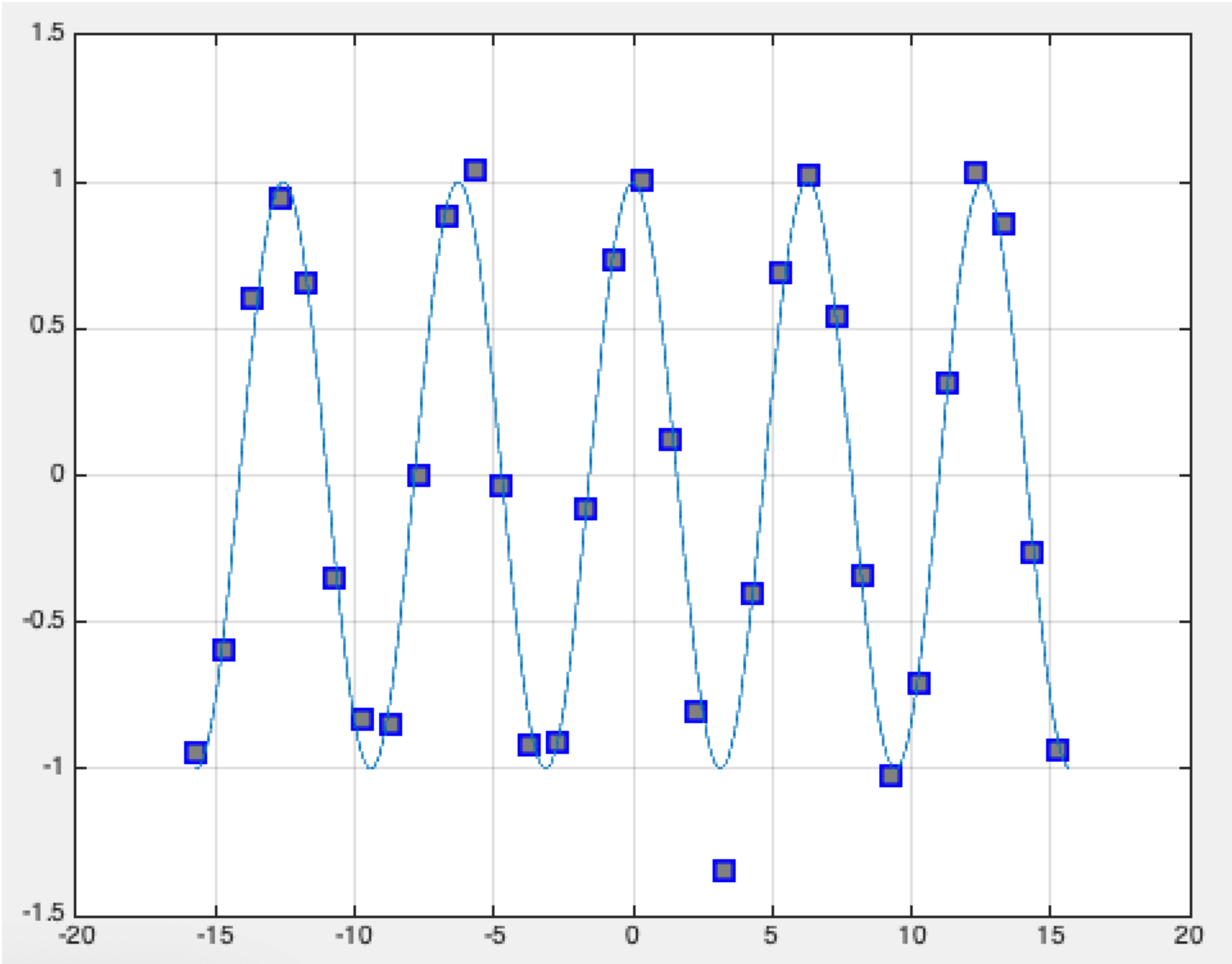
Modèle 1 : $x_i = a + e_i$ avec $e_i \sim \mathcal{N}(0, \sigma^2)$



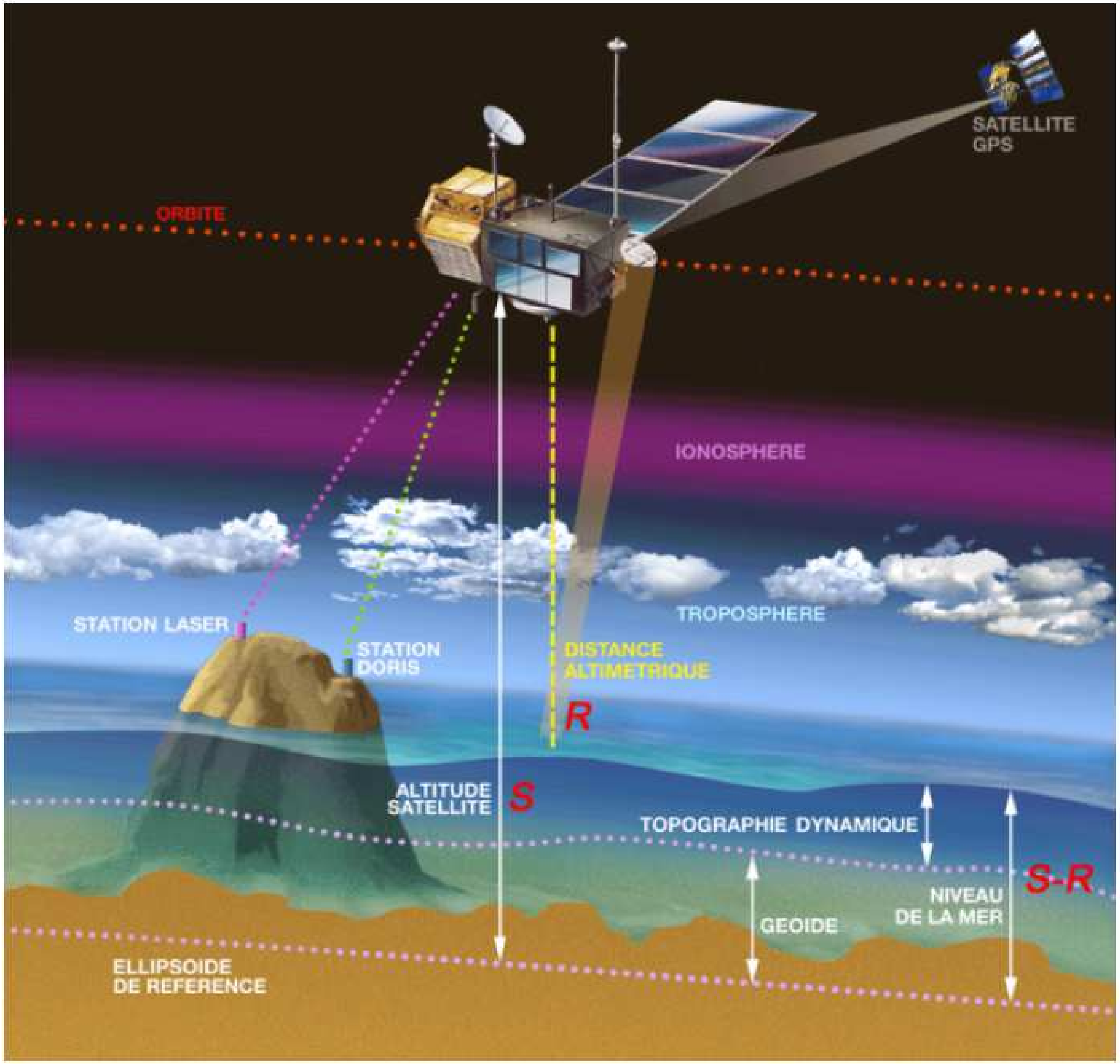
Modèle 2 : $x_i = ai + b + e_i$ avec $e_i \sim \mathcal{N}(0, \sigma^2)$



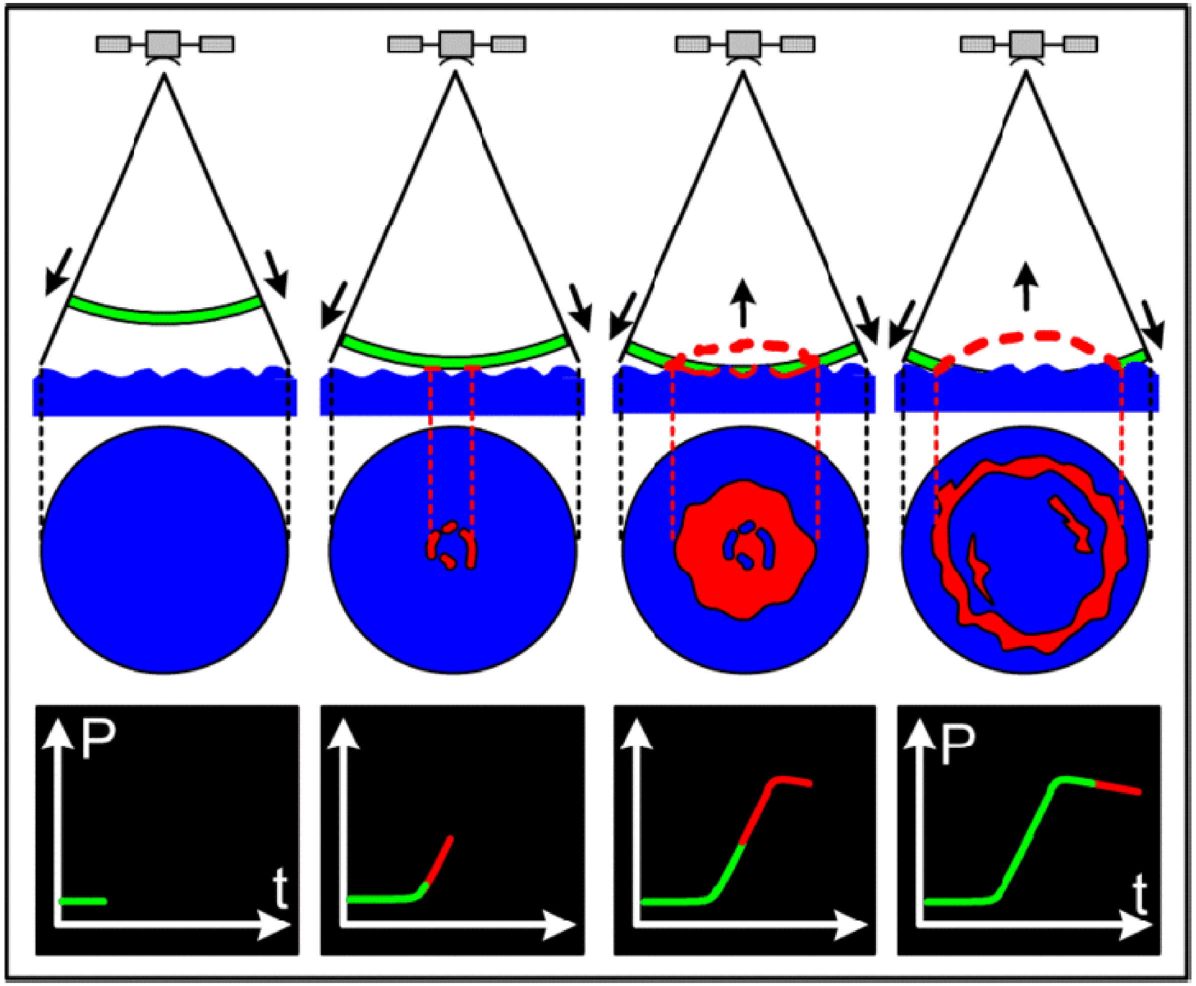
Modèle 3 : $x_i = a\cos(i\phi) + e_i$ avec $e_i \sim \mathcal{N}(0, \sigma^2)$



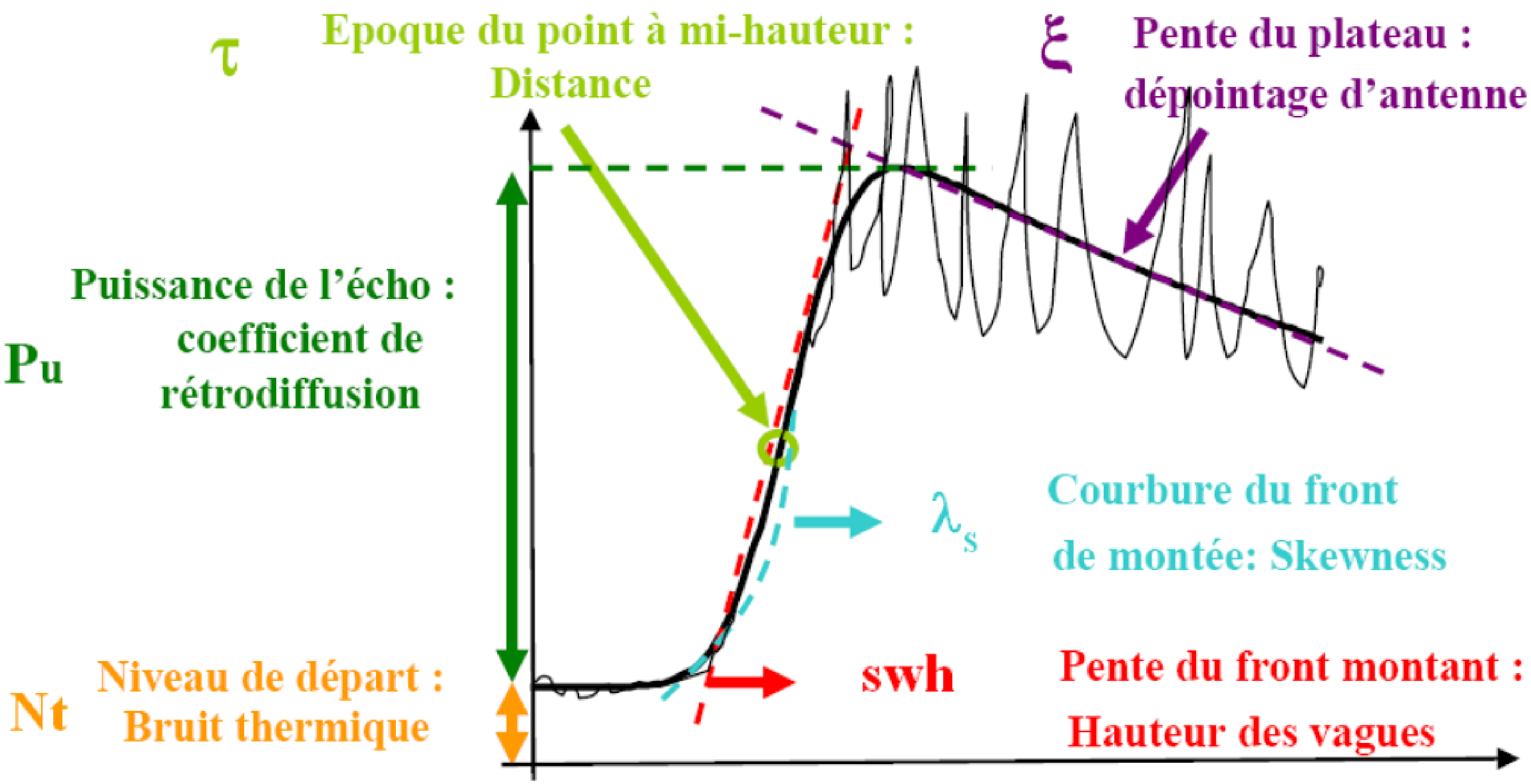
Altimétrie


SATELLITE
GPS
ORBITE
IONOSPHERE
TROPOSPHERE
STATION LASER
STATION
DORIS
DISTANCE
ALTIMETRIQUE
R
ALTITUDE
SATELLITE
S
TOPOGRAPHIE DYNAMIQUE
S-R
NIVEAU
DE LA MER
GEOIDE
ELLIPSOIDE
DE REFERENCE

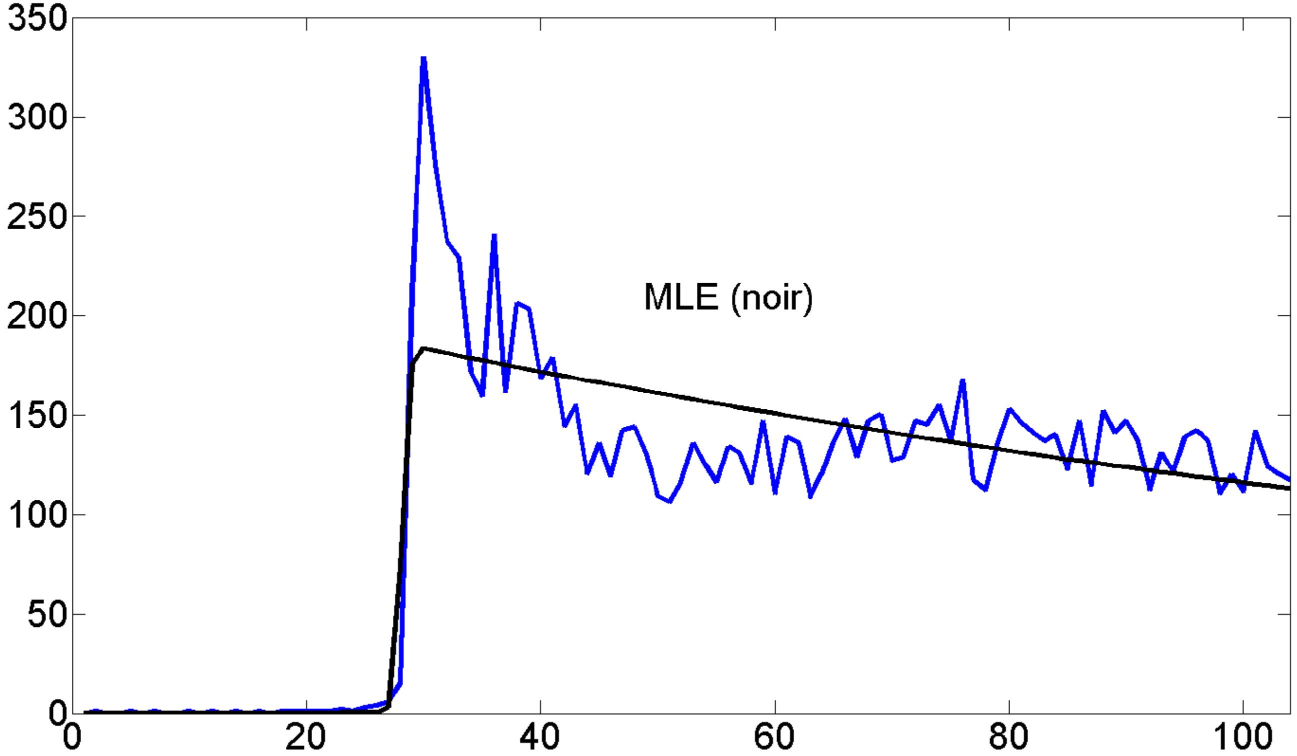
Formation de l’écho altimétrique



Modèle de Brown



Modèle de Brown



Qualités d'un estimateur

- $\theta \in \mathbb{R}$
  - Biais (erreur systématique) : $b_n(\theta) = E\left(\widehat{\theta}_n\right) - \theta$
  - Variance

$$v_n(\theta) = E\left[\left(\widehat{\theta}_n - E\left(\widehat{\theta}_n\right)\right)^2\right] = E\left[\widehat{\theta}_n^2\right] - E\left(\widehat{\theta}_n\right)^2$$

  - Erreur quadratique moyenne (précision)

$$e_n(\theta) = E\left[\left(\widehat{\theta}_n - \theta\right)^2\right] = v_n(\theta) + b_n^2(\theta)$$

CS de convergence : $\widehat{\theta}_n$ est un estimateur convergent si $\lim\limits_{n \to +\infty} b_n(\theta) = \lim\limits_{n \to +\infty} v_n(\theta) = 0$

Qualités d'un estimateur

- $\boldsymbol{\theta} \in \mathbb{R}^p$
  - Biais

$$b_n(\boldsymbol{\theta}) = E\left(\widehat{\boldsymbol{\theta}}_n\right) - \boldsymbol{\theta} \in \mathbb{R}^p$$

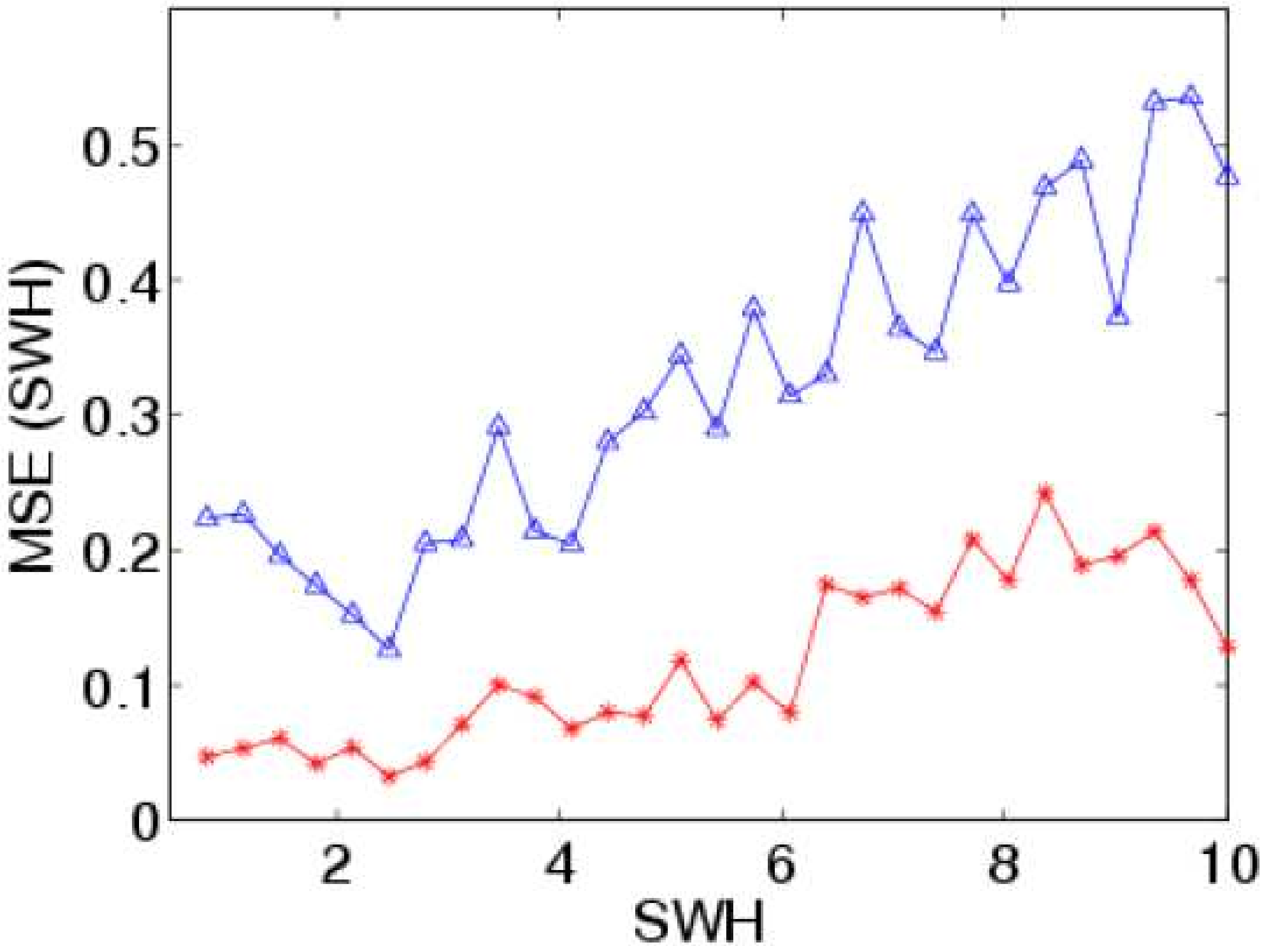
  - Matrice de covariance

$$E\left[\left(\widehat{\boldsymbol{\theta}}_n - E\left(\widehat{\boldsymbol{\theta}}_n\right)\right)\left(\widehat{\boldsymbol{\theta}}_n - E\left(\widehat{\boldsymbol{\theta}}_n\right)\right)^T\right]$$

Exemples

- Exemple 1 : $X_i \sim \mathcal{N}(m, \sigma^2)$, $\theta = m$ et σ^2 connue
  - Moyenne empirique
$$\widehat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$
  - Autre estimateur
$$\widetilde{\theta}_n = \frac{2}{n(n+1)}\sum_{i=1}^{n} iX_i$$
- Exemple 2 : $X_i \sim \mathcal{N}(m, \sigma^2)$, $\theta = \sigma^2$, m connue ou inconnue
- Exemple 3 : $X_i \sim \mathcal{N}(m, \sigma^2)$, $\boldsymbol{\theta} = (m, \sigma^2)^T$

Mean Square Errors



Plan du cours

- Chapitre 1 : Convergence et théorèmes limites
- Chapitre 2 : Estimation
  - Modèle statistique, qualités d'un estimateur, exemples
  - Inégalité de Cramér Rao
  - Maximum de vraisemblance
  - Méthode des moments
  - Estimation Bayésienne
  - Intervalles de confiance
- Chapitre 3 : Tests Statistiques

Inégalité de Cramér Rao

- Vraisemblance

$$L(x_1, \ldots, x_n; \theta) = \begin{cases} X_i \text{ va discrète : } & P[X_1 = x_1, \ldots, X_n = x_n; \theta] \\ X_i \text{ va continue : } & p(x_1, \ldots, x_n; \theta) \end{cases}$$

- Inégalité pour $\boldsymbol{\theta} \in \mathbb{R}$
  - Définition

$$\text{Var}\left(\widehat{\theta}_n\right) \geq \frac{\left[1 + b_n'(\theta)\right]^2}{-E\left[\frac{\partial^2 \ln L(X_1, \ldots, X_n; \theta)}{\partial \theta^2}\right]} = \text{BCR}(\theta)$$

BCR(θ) est appelée Borne de Cramér Rao de θ

  - Hypothèses
    Log-vraisemblance deux fois dérivable et support de la loi indépendant de θ (contre-exemple : loi $\mathcal{U}[0, \theta]$)

Inégalité de Cramér Rao

- Estimateur Efficace : estimateur sans biais tel que $\text{Var}\left(\widehat{\theta}_n\right) = \text{BCR}(\theta)$ (Il est unique !)
  Exemple : $X_i \sim \mathcal{N}(m, \sigma^2)$, $\theta = m$ et σ^2 connue
- Cas où $(X_1, ..., X_n)$ est un échantillon

$$\text{Var}\left(\widehat{\theta}_n\right) \geq \frac{\left[1 + b_n'(\theta)\right]^2}{-nE\left[\frac{\partial^2 \ln L(X_1;\theta)}{\partial \theta^2}\right]} = \text{BCR}(\theta)$$

Cas multivarié

- Inégalité pour un estimateur non biaisé de $\boldsymbol{\theta} \in \mathbb{R}^p$

$$\text{Cov}\left(\widehat{\boldsymbol{\theta}}\right) \geq I_n^{-1}(\boldsymbol{\theta})$$

avec $I_{ij} = E\left[-\frac{\partial^2 \ln L(X_1,...,X_n;\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right]$ pour $i, j = 1, ..., p$ et
$A \geq B$ signifie $A - B$ matrice semi définie positive

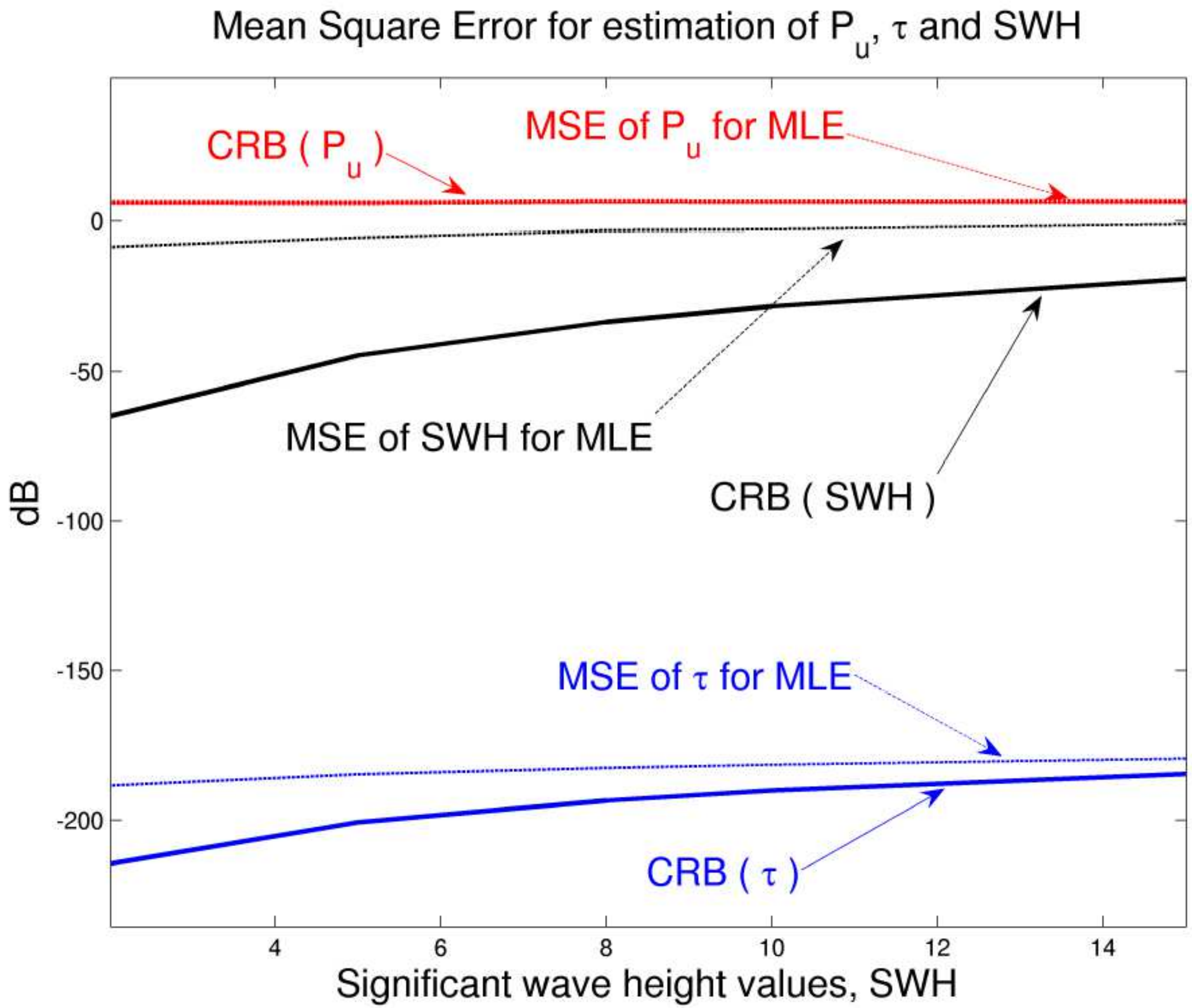
$$x^T(A - B)x \geq 0, \quad \forall x \in \mathbb{R}^p$$

On en déduit

$$\text{Var}\left(\widehat{\theta_i}\right) \geq \left[I_n^{-1}(\boldsymbol{\theta})\right]_{ii}$$

Exemple : $X_i \sim \mathcal{N}(m, \sigma^2)$, $\boldsymbol{\theta} = (m, \sigma^2)^T$.

Borne de Cramér-Rao



Plan du cours

- Chapitre 1 : Convergence et théorèmes limites
- Chapitre 2 : Estimation
  - Modèle statistique, qualités d'un estimateur, exemples
  - Inégalité de Cramér Rao
  - Maximum de vraisemblance
  - Méthode des moments
  - Estimation Bayésienne
  - Intervalles de confiance
- Chapitre 3 : Tests Statistiques

Méthode du Maximum de Vraisemblance

- Définition

$$\widehat{\boldsymbol{\theta}}_{\text{MV}} = \arg\max_{\theta} L(X_1, ..., X_n; \boldsymbol{\theta})$$

- Recherche du maximum pour $\theta \in \mathbb{R}$
  Si $L(X_1, ..., X_n; \theta)$ est régulière, on résoud

$$\frac{\partial L(X_1, ..., X_n; \theta)}{\partial \theta} = 0 \text{ ou } \frac{\partial \ln L(X_1, ..., X_n; \theta)}{\partial \theta} = 0$$

  et on vérifie qu'on a bien un maximum en faisant un tableau de variations ou en étudiant

$$\frac{\partial^2 \ln L(X_1, ..., X_n; \widehat{\theta}_{\text{MV}})}{\partial \theta^2} < 0$$

Régularité

On dit qu'une variable aléatoire X de densité de probabilité $f(x;\theta)$ est régulière si (voir livre de Lehmann, Theory of Point Estimation)

- Le support de la densité f, i.e., $\{x|f(x;\theta)>0\}$, est indépendant de θ
- $f(x;\theta)$ est au moins trois fois dérivable par rapport à θ
- La vraie valeur de θ appartient à un ensemble compact Θ.

Dans ce cas, la recherche de l'estimateur du maximum de vraisemblance peut se faire en cherchant les racines de

$$\frac{\partial L(X_1, ..., X_n;\theta)}{\partial \theta} = 0 \text{ ou } \frac{\partial \ln L(X_1, ..., X_n;\theta)}{\partial \theta} = 0$$

Méthode du Maximum de Vraisemblance

- Recherche du maximum pour $\theta \in \mathbb{R}^p$

$$\frac{\partial L(X_1, ..., X_n; \theta)}{\partial \theta_i} = 0 \text{ ou } \frac{\partial \ln L(X_1, ..., X_n; \theta)}{\partial \theta_i} = 0$$

pour $i = 1, ..., p$

- Exemples
  - Exemple 1 : $X_i \sim \mathcal{P}(\lambda)$, $\theta = \lambda$
  - Exemple 2 : $X_i \sim \mathcal{N}(m, \sigma^2)$, $\boldsymbol{\theta} = (m, \sigma^2)^T$

Propriétés

- Estimateur asymptotiquement non biaisé

$$\lim_{n \to +\infty} E\left[\widehat{\boldsymbol{\theta}}_{\text{MV}}\right] - \boldsymbol{\theta} = 0$$

- Estimateur convergent
- Estimateur asymptotiquement efficace

$$\lim_{n \to +\infty} \frac{\text{Var}\left(\widehat{\theta_i}\right)}{\left[I_n^{-1}(\boldsymbol{\theta})\right]_{ii}} = 1$$

- Normalité Asymptotique

Propriétés

- Invariance Fonctionnelle
  Si $\boldsymbol{\mu} = h(\boldsymbol{\theta})$, où h est une fonction bijective d'un ouvert $O \subset \mathbb{R}^p$ dans un ouvert $V \subset \mathbb{R}^p$, alors

$$\widehat{\boldsymbol{\mu}}_{\text{MV}} = h\left(\widehat{\boldsymbol{\theta}}_{\text{MV}}\right)$$

- Conclusions
  L'estimateur $\widehat{\boldsymbol{\theta}}_{\text{MV}}$ possède beaucoup de bonnes propriétés asymptotiques mais peut être difficile à étudier car il est la solution d'un problème d'optimisation.

Remarques sur la convergence

Théorème (voir livre de Lehmann, Theory of Point Estimation) : soient $X_1, ..., X_n$ des variables aléatoires iid de même densité $f(x_i;\theta)$ avec les hypothèses suivantes

- θ appartient à un ouvert $\Theta \in \mathbb{R}$
- le paramètre θ est identifiable, i.e., deux valeurs différentes de θ donnent des densités $f(x_i;\theta)$ différentes
- la log-vraisemblance $l(\theta)$ est dérivable par rapport à θ
- le support de la densité f ne dépend pas de θ

alors l'équation $l'(\theta) = 0$ admet une solution qui converge en probabilité vers θ_0 (pas nécessairement $\hat{\theta}_{\text{MV}}$). Donc s'il y a une unique solution de $l'(\theta) = 0$ et que cette solution est un maximum local de la vraisemblance, alors cette solution est l'estimateur du maximum de vraisemblance de θ et cet estimateur est convergent.

Exemple d'estimateur $\widehat{\theta}_{\text{MV}}$ non convergent (avec plusieurs maxima locaux de $l'(\theta) = 0$)

$$f(x_i;\theta) = \frac{1}{2}\mathcal{N}(0,1) + \frac{1}{2}\mathcal{N}\left(\theta, [\exp(-1/\theta^2)]^2\right)$$

Plan du cours

- Chapitre 1 : Convergence et théorèmes limites
- Chapitre 2 : Estimation
  - Modèle statistique, qualités d'un estimateur, exemples
  - Inégalité de Cramér Rao
  - Maximum de vraisemblance
  - Méthode des moments
  - Estimation Bayésienne
  - Intervalles de confiance
- Chapitre 3 : Tests Statistiques

Méthode des moments

- Définition : supposons que $X_1, ..., X_n$ ont la même loi de paramètre inconnu $\boldsymbol{\theta} \in \mathbb{R}^p$. En général, le vecteur paramètre à estimer $\boldsymbol{\theta}$ est lié aux premiers moments de la loi commune des va X_i par une relation notée

$$\boldsymbol{\theta} = h(m_1, ..., m_q)$$

avec $m_k = E[X_i^k]$ et $q \geq p$. Un estimateur des moments de $\boldsymbol{\theta}$ est défini par

$$\widehat{\boldsymbol{\theta}}_{\text{Mo}} = h(\widehat{m}_1, ..., \widehat{m}_q) \text{ avec } \widehat{m}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

- Exemple : $X_i \sim \mathcal{N}(m, \sigma^2)$, $\boldsymbol{\theta} = (m, \sigma^2)^T$

Propriétés

- Estimateur convergent
- Normalité Asymptotique
- Conclusion : peu de propriétés mais cet estimateur est généralement facile à étudier.

Plan du cours

- Chapitre 1 : Convergence et théorèmes limites
- Chapitre 2 : Estimation
  - Modèle statistique, qualités d'un estimateur, exemples
  - Inégalité de Cramér Rao
  - Maximum de vraisemblance
  - Méthode des moments
  - Estimation Bayésienne
  - Intervalles de confiance
- Chapitre 3 : Tests Statistiques

Estimation Bayésienne

- Principe : l'estimation Bayésienne consiste à estimer un vecteur paramètre inconnu $\boldsymbol{\theta} \in \mathbb{R}^p$ à l'aide de la vraisemblance de $X_1, ..., X_n$ (paramétrée par $\boldsymbol{\theta}$) et d'une loi *a priori* $p(\boldsymbol{\theta})$. Pour cela, on minimise une fonction de coût $c\left(\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}}\right)$ qui représente l'erreur entre $\boldsymbol{\theta}$ et $\widehat{\boldsymbol{\theta}}$.
- Estimateur MMSE : l'estimateur qui minimise l'erreur quadratique moyenne (mean square error (MSE)) $c\left(\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}}\right) = E\left[\left(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\right)^T \left(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\right)\right]$ est

$$\widehat{\boldsymbol{\theta}}_{\text{MMSE}} = E\left(\boldsymbol{\theta} | X_1, ..., X_n\right)$$

  - Remarque : $p\left(\boldsymbol{\theta} | x_1, ..., x_n\right)$ est la loi *a posteriori* de $\boldsymbol{\theta}$
  - Preuve : voir cours

Estimation Bayésienne

- Estimateur MAP : l'estimateur du maximum a posteriori (MAP) est défini par

$$\widehat{\boldsymbol{\theta}}_{\text{MAP}} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|X_1, ..., X_n)$$

Cet estimateur minimise la fonction de coût $E\left[c\left(\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}}\right)\right]$ avec

$$c\left(\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}}\right) = \begin{cases} 1 \text{ si } \left\|\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\right\| > \Delta \\ 0 \text{ si } \left\|\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\right\| < \Delta \end{cases}$$

avec Δ arbitrairement petit.

- Preuve : voir livre de H. Van Trees

Exemple

- Vraisemblance

$$X_i \sim \mathcal{N}(\theta, \sigma^2)$$

- Loi a priori

$$\theta \sim \mathcal{N}(\mu, \nu^2)$$

- Loi a posteriori

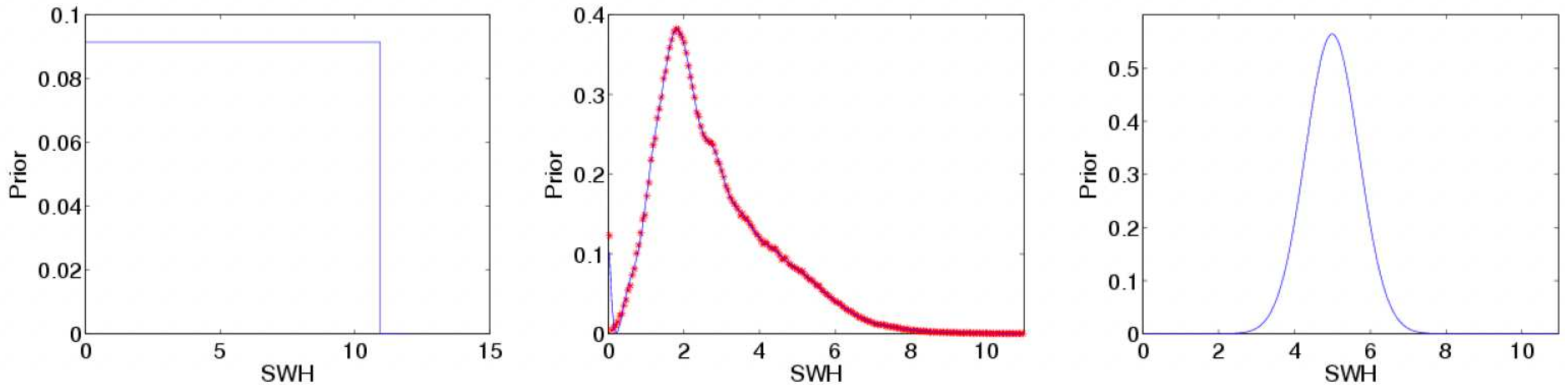
$$\theta | X_1, ..., X_n \sim \mathcal{N}\left(m_p, \sigma_p^2\right)$$

- Estimateurs Bayésiens

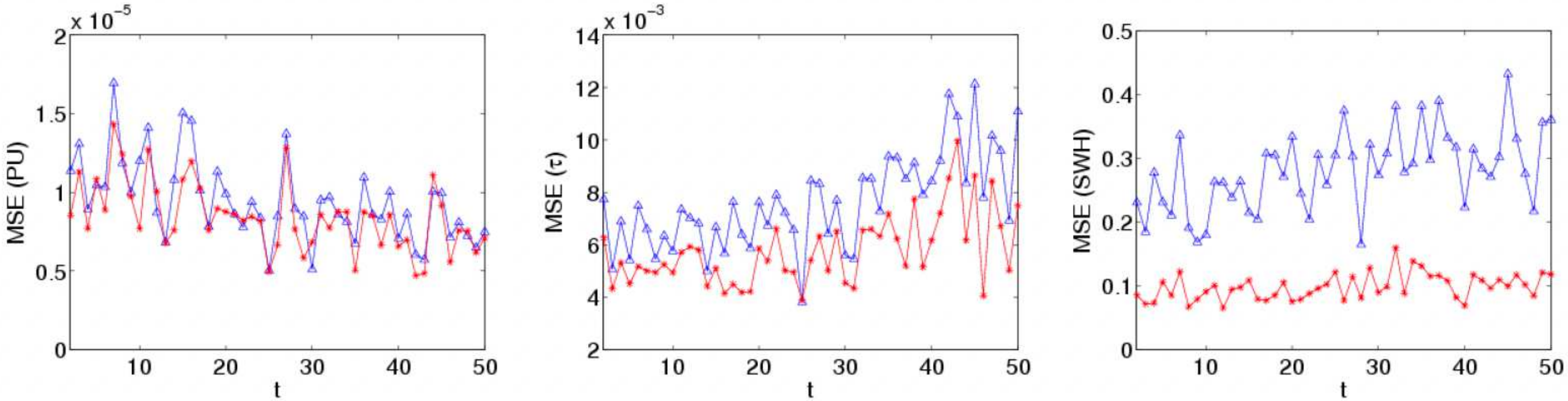
$$\widehat{\theta}_{\text{MAP}} = \widehat{\theta}_{\text{MMSE}} = m_p = \overline{X}\left(\frac{n\nu^2}{n\nu^2 + \sigma^2}\right) + \mu\left(\frac{\sigma^2}{\sigma^2 + n\nu^2}\right)$$

Avec ou sans prior ?

Examples for the Significant Wave Height (SWH)



Dynamic priors



Plan du cours

- Chapitre 1 : Convergence et théorèmes limites
- Chapitre 2 : Estimation
  - Modèle statistique, qualités d'un estimateur, exemples
  - Inégalité de Cramér Rao
  - Maximum de vraisemblance
  - Méthode des moments
  - Estimation Bayésienne
  - Intervalles de confiance
- Chapitre 3 : Tests Statistiques

Intervalles de confiance

- Principe : un intervalle de confiance $[a, b]$ pour le paramètre $\theta \in \mathbb{R}$ est un intervalle tel que $P[a < \theta < b] = \alpha$, où α est le paramètre de confiance (en général $\alpha = 0.99$ ou $\alpha = 0.95$).
- Détermination pratique de l'intervalle : on cherche un estimateur de θ noté $\widehat{\theta}$ (par la méthode des moments, du maximum de vraisemblance, ...), on en déduit une statistique $T(X_1, ..., X_n)$ qui dépend de θ de loi connue, on cherche $c(\theta)$ et $d(\theta)$ tels que

$$P[c(\theta) < T(X_1, ..., X_n) < d(\theta)] = \alpha$$

On en déduit l'intervalle $[a, b]$.

Exemples

- Exemple 1 : $X_i \sim \mathcal{N}(m, \sigma^2)$, m inconnue, σ^2 connue.

$$T = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i - m}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

- Exemple 2 : $X_i \sim \mathcal{N}(m, \sigma^2)$, IC pour m, σ^2 inconnue.

$$T \sim \mathcal{N}(0,1) \quad \text{et} \quad U = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \sim \chi^2_{n-1}$$

donc

$$\frac{T}{\sqrt{\frac{U}{n-1}}} \sim t_{n-1}$$

suit une loi de Student à $n-1$ degrés de liberté.

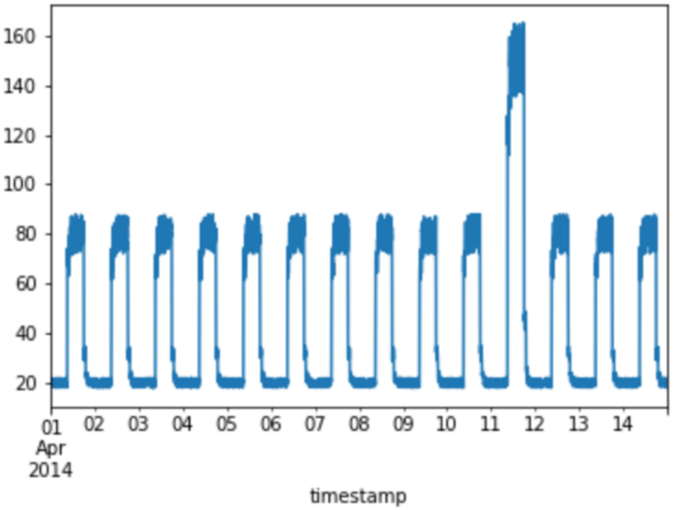
Que faut-il savoir ?

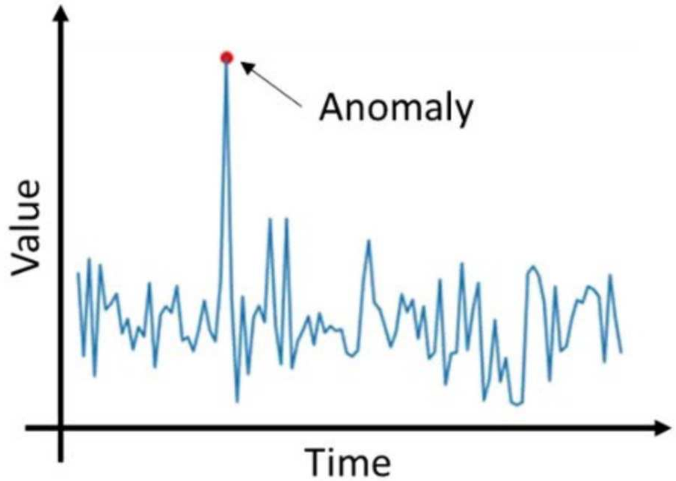
- Notions de biais, variance et convergence d'un estimateur
- Calcul d'une borne de Cramér-Rao et notion d'efficacité
- Détermination de l'estimateur du maximum de vraisemblance (MV)
- Propriétés de l'estimateur MV
- Principe et application de la méthode des moments
- Principe et application de l'estimation Bayésienne
- Détermination des intervalles de confiance

Plan du cours

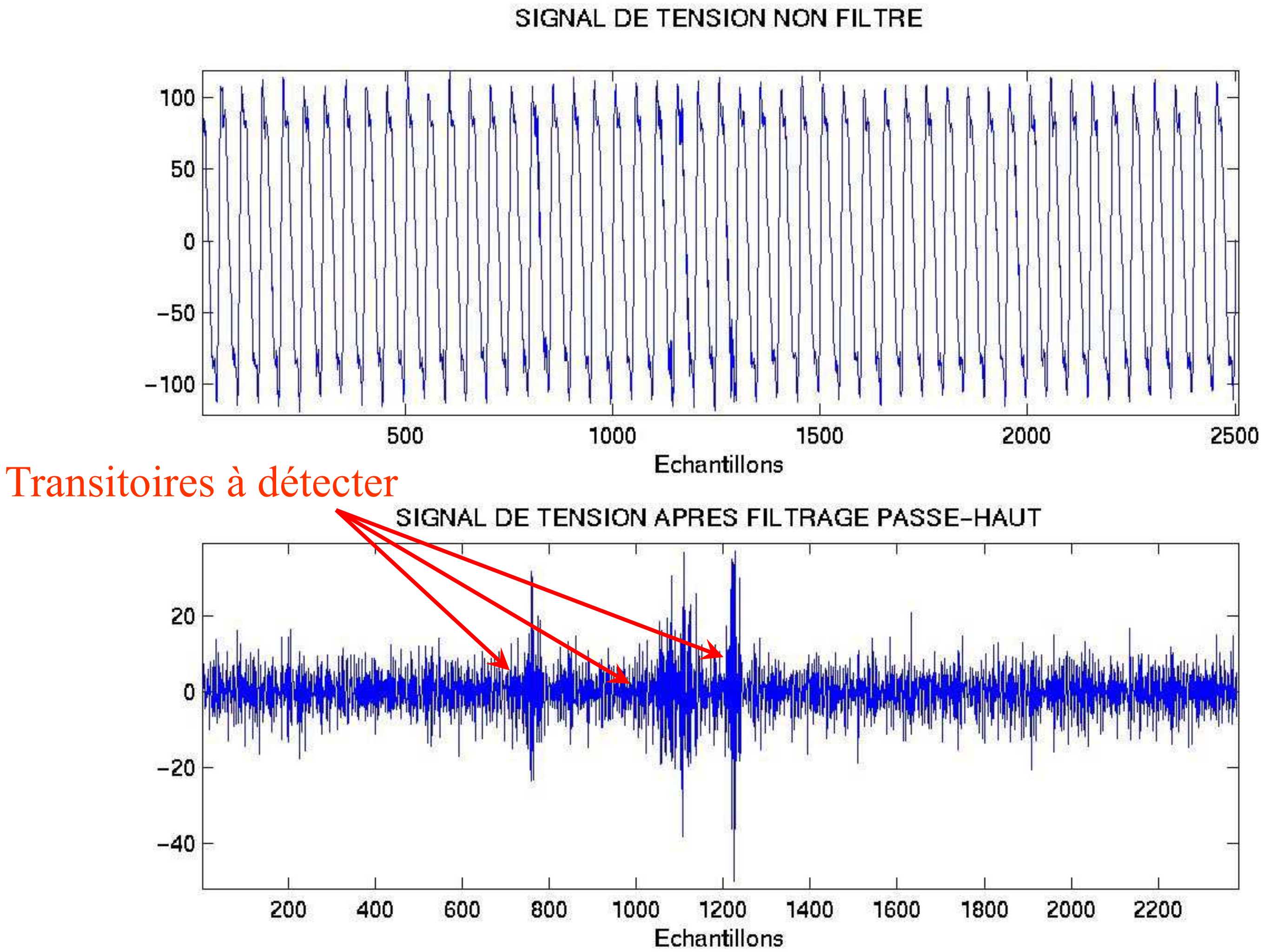
- Chapitre 1 : Convergence et théorèmes limites
- Chapitre 2 : Estimation
- Chapitre 3 : Tests Statistiques
  - Généralités, exemple
  - Courbes COR
  - Théorème de Neyman Pearson
  - Autres tests paramétriques
  - Test du χ^2
  - Test de Kolmogorov

Problème


160
140
120
100
80
60
40
20
01
Apr
2014
02
03
04
05
06
07
08
09
10
11
12
13
14
timestamp


Anomaly
Value
Time

Problème


SIGNAL DE TENSION NON FILTRE
100
50
0
−50
−100
500
1000
1500
2000
2500
Echantillons
Transitoires à détecter
SIGNAL DE TENSION APRES FILTRAGE PASSE-HAUT
20
0
−20
−40
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
Echantillons

Généralités

- Principe : un test statistique est un mécanisme qui permet de décider entre plusieurs hypothèses $H_0, H_1, ...$ à partir de n observations $x_1, ..., x_n$. On se limitera dans ce cours à deux hypothèses H_0 et H_1. Effectuer un test, c'est déterminer une statistique de test $T(X_1, ..., X_n)$ et un ensemble Δ tel que

$$\begin{aligned} &\mathcal{H}_0 \text{ rejetée si } T(X_1, ..., X_n) \in \Delta \\ &\mathcal{H}_0 \text{ acceptée si } T(X_1, ..., X_n) \notin \Delta. \end{aligned} \quad (1)$$

- Vocabulaire
  - H_0 est l'hypothèse nulle
  - H_1 est l'hypothèse alternative
  - $\{(x_1, ..., x_n) | T(x_1, ..., x_n) \in \Delta\}$: région critique

Définitions

- Tests paramétriques et non paramétriques
- Hypothèses simples et hypothèses composites
- Risque de première espèce = probabilité de fausse alarme

$$\alpha = \text{PFA} = P[\text{Rejeter } H_0 | H_0 \text{ vraie}]$$

- Risque de seconde espèce = probabilité de non-détection

$$\beta = \text{PND} = P[\text{Rejeter } H_1 | H_1 \text{ vraie}]$$

- Puissance du test = probabilité de détection : $\pi = 1 - \beta$

Exemple

$$X_i \sim \mathcal{N}(m, \sigma^2),\ \sigma^2 \text{ connue}$$

- Hypothèses

$$H_0 : m = m_0,\ H_1 : m = m_1 > m_0$$

- Stratégie du test

$$\text{Rejet de } H_0 \text{ si } \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i > S_\alpha$$

- Problèmes
  Déterminer le seuil S_α, le risque β et la puissance du test π.

Plan du cours

- Chapitre 1 : Convergence et théorèmes limites
- Chapitre 2 : Estimation
- Chapitre 3 : Tests Statistiques
  - Généralités, exemple
  - Courbes COR, p-valeur
  - Théorème de Neyman Pearson
  - Autres tests paramétriques
  - Test du χ^2
  - Test de Kolmogorov

Courbes COR

- Caractéristiques opérationnelles du récepteur

$$\text{PD} = h\left(\text{PFA}\right)$$

- Exemple : $X_i \sim \mathcal{N}(m, \sigma^2), \ \sigma^2$ connue

$$H_0 : m = m_0, \ H_1 : m = m_1 > m_0$$

  - Probabilité de fausse alarme

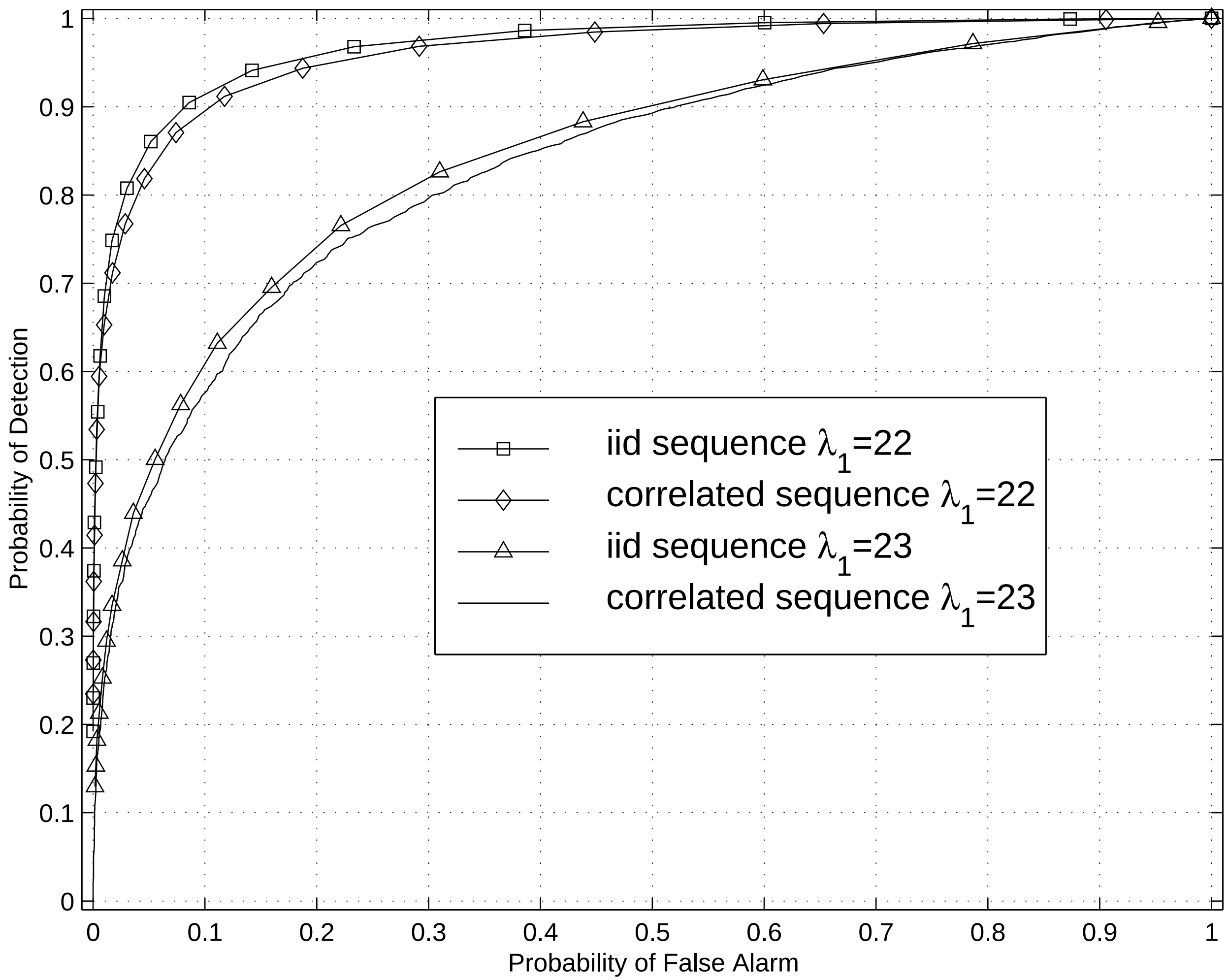
$$\alpha = 1 - F\left(\frac{S_\alpha - m_0}{\frac{\sigma}{\sqrt{n}}}\right) \Leftrightarrow S_\alpha = m_0 + \frac{\sigma}{\sqrt{n}} F^{-1}(1-\alpha)$$

  - Probabilité de détection

$$\text{PD} = \pi = 1 - F\left(\frac{S_\alpha - m_1}{\frac{\sigma}{\sqrt{n}}}\right)$$

Représentation graphique

ROC's for iid and correlated sequences



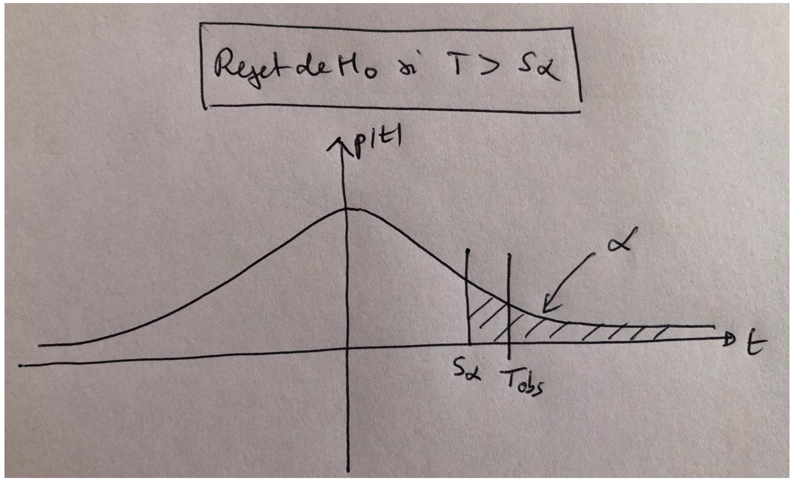
p-valeur d'un test

- Définition

$$p(\boldsymbol{x}) = \inf\{\alpha \in]0,1[\mid \boldsymbol{x} \in \mathcal{R}_\alpha\}$$

où $\mathcal{R}_\alpha$ est la zone de rejet pour α fixé et $\boldsymbol{x} = (x_1, ..., x_n)$. C'est la plus petite valeur de α pour laquelle on rejette H_0.

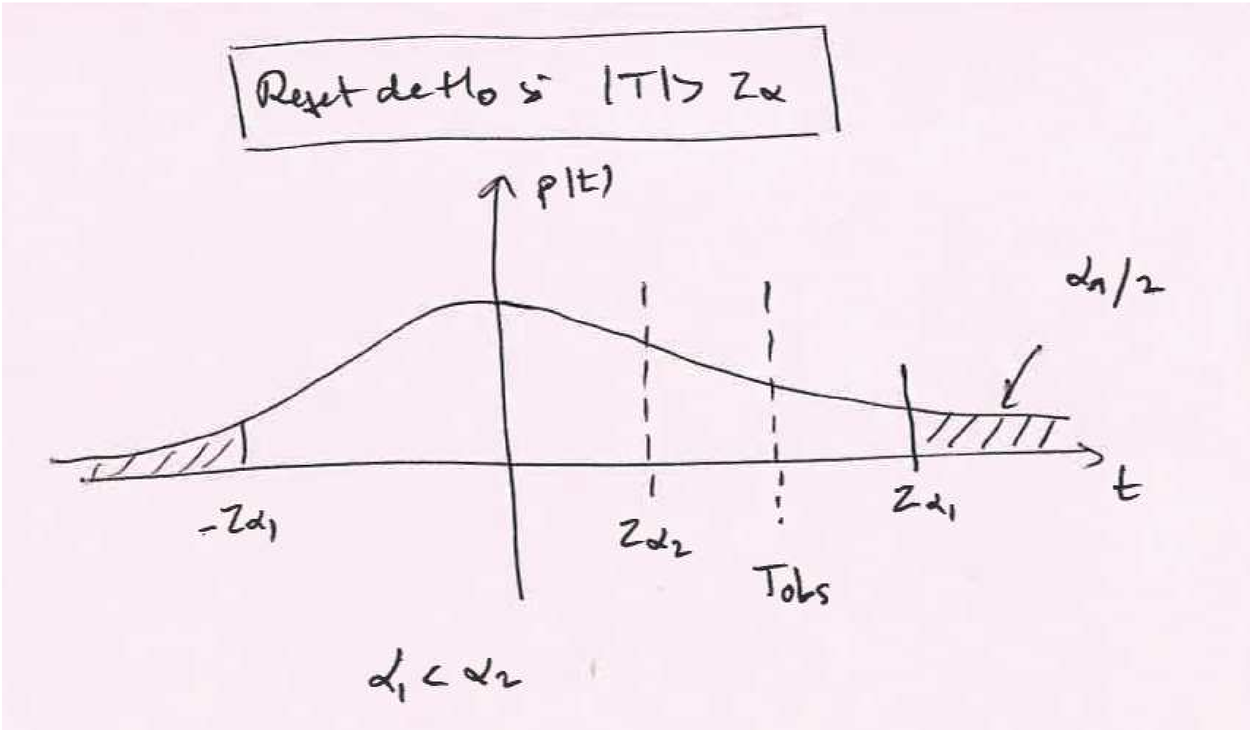
- Calcul



  - Si $\alpha = 0$, on accepte toujours H_0 donc $S_0 = +\infty$
  - Si $\alpha = 1$, on rejette toujours H_0 donc $S_1 = -\infty$
  - Plus petite valeur de α pour laquelle on rejette H_0

$$\alpha^* = 1 - F(T_{\text{obs}}).$$

Autre exemple



- Si $\alpha = 0$, on accepte toujours H_0 donc $z_0 = 0$
- Si $\alpha = 1$, on rejette toujours H_0 donc $z_1 = +\infty$
- Plus petite valeur de α pour laquelle on rejette H_0

$$\frac{\alpha^*}{2} = 1 - F(|T_{\text{obs}}|) \Leftrightarrow \alpha^* = 2[1 - F(|T_{\text{obs}})|].$$

Plan du cours

Théorème de Neyman-Pearson

Test paramétrique à hypothèses simples

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \text{ et } H_1 : \boldsymbol{\theta} = \boldsymbol{\theta}_1 \tag{2}$$

- Variables aléatoires continues
  - Théorème : à α fixé, le test qui minimise β (ou maximise π) est défini par

$$\text{Rejet de } H_0 \text{ si } \frac{L(x_1, ..., x_n | H_1)}{L(x_1, ..., x_n | H_0)} > S_\alpha$$

  - Remarque : $L(x_1, ..., x_n | H_i) = f(x_1, ..., x_n | \boldsymbol{\theta}_i)$
  - Exemple : $X_i \sim \mathcal{N}(m, \sigma^2), \; \sigma^2$ connue

$$H_0 : m = m_0, \; H_1 : m = m_1 > m_0$$

Résumé

Effectuer un test de Neyman-Pearson, c'est

- 1) Déterminer la statistique et la région critique du test
- 2) Déterminer la relation entre le seuil S_α et le risque α
- 3) Calculer le risque β et la puissance π du test (ou la courbe COR)
- 4) Application numérique : on accepte ou rejette l'hypothèse H_0 en précisant le risque α donné

Théorème de Neyman-Pearson

- Variables aléatoires discrètes
  - Théorème : parmi tous les tests de risque de première espèce $\leq \alpha$ fixé, le test de puissance maximale est défini par

$$\text{Rejet de } H_0 \text{ si } \frac{L(x_1, ..., x_n | H_1)}{L(x_1, ..., x_n | H_0)} > S_\alpha$$

  - Remarque :

$$L(x_1, ..., x_n | H_i) = P\left[X_1 = x_1, ..., X_n = x_n | \boldsymbol{\theta}_i\right]$$

  - Exemple : $X_i \sim \mathcal{P}(\lambda)$, $H_0 : \lambda = \lambda_0, \; H_1 : \lambda = \lambda_1 > \lambda_0$
  - Loi asymptotique : quand n est suffisamment grand, utilisation du théorème de la limite centrale

Plan du cours

- Chapitre 1 : Convergence et théorèmes limites
- Chapitre 2 : Estimation
- Chapitre 3 : Tests Statistiques
  - Généralités, exemple
  - Courbes COR
  - Théorème de Neyman Pearson
  - Autres tests paramétriques
  - Test du χ^2
  - Test de Kolmogorov

Test du rapport de vraisemblance généralisé

Test paramétrique à hypothèses composites

$$H_0 : \boldsymbol{\theta} \in \Theta_0 \text{ et } H_1 : \boldsymbol{\theta} \in \Theta_1 \tag{3}$$

- Définition (Test GLR)

$$\text{Rejet de } H_0 \text{ si } \frac{L\left(x_1, ..., x_n | \widehat{\boldsymbol{\theta}}_1^{\text{MV}}\right)}{L\left(x_1, ..., x_n | \widehat{\boldsymbol{\theta}}_0^{\text{MV}}\right)} > S_\alpha$$

où $\widehat{\boldsymbol{\theta}}_0^{\text{MV}}$ et $\widehat{\boldsymbol{\theta}}_1^{\text{MV}}$ sont les estimateurs du maximum de vraisemblance de $\boldsymbol{\theta}$ sous les hypothèses H_0 et H_1.

- Remarque

$$L\left(x_1, ..., x_n | \widehat{\boldsymbol{\theta}}_i^{\text{MV}}\right) = \sup_{\boldsymbol{\theta} \in \Theta_i} L\left(x_1, ..., x_n | \boldsymbol{\theta}\right)$$

Tests sur l'espérance d'un échantillon gaussien

Soit $(X_1, ..., X_n)$ un échantillon gaussien de loi $\mathcal{N}(m, \sigma^2)$ avec une variance σ^2 connue. On considère les hypothèses

$$H_0 : m = m_0 \text{ et } H_1 : m = m_1 \text{ avec } m_1 > m_0 \tag{4}$$

- Définition du test

$$\text{Rejet de } H_0 \text{ si } T = \sqrt{n}\left(\frac{\overline{X} - m_0}{\sigma}\right) > S_\alpha \quad \text{avec} \quad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- Remarques
  - T suit une loi normale sous H_0
  - Application directe de Neyman-Pearson
  - Généralisation immédiate à $m_1 < m_0$ ou à $m_1 \neq m_0$

Tests sur l'espérance d'un échantillon gaussien

Soit $(X_1, ..., X_n)$ un échantillon gaussien de loi $\mathcal{N}(m, \sigma^2)$ avec une variance σ^2 inconnue et les hypothèses

$$H_0 : m = m_0 \text{ et } H_1 : m = m_1 \text{ avec } m_1 > m_0 \tag{5}$$

- Définition du test

  Rejet de H_0 si $T = \sqrt{n}\left(\frac{\overline{X} - m_0}{S_n}\right) > S_\alpha$ avec $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$

- Remarques
  - Si on pose $U = \sqrt{n}\left(\frac{\overline{X}-m_0}{\sigma}\right)$ et $V = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2$, on a $T = \frac{U}{\sqrt{\frac{V}{n-1}}}$ qui suit donc une loi de Student à $n-1$ ddl sous H_0
  - Généralisation immédiate à $m_1 < m_0$ ou à $m_1 \neq m_0$

Tests sur la variance d'un échantillon gaussien

Soit $(X_1, ..., X_n)$ un échantillon gaussien de loi $\mathcal{N}(m, \sigma^2)$ avec une moyenne m connue et les hypothèses

$$H_0 : \sigma^2 = \sigma_0^2 \text{ et } H_1 : \sigma^2 = \sigma_1^2 \text{ avec } \sigma_1^2 > \sigma_0^2 \qquad (6)$$

- Définition du test

$$\text{Rejet de } H_0 \text{ si } T = \frac{1}{\sigma_0^2} \sum_{i=1}^{n} (X_i - m)^2 > S_\alpha$$

- Remarques
  - La loi de T sous H_0 est une loi du χ_n^2, ce qui permet de déterminer S_α en fonction de α.
  - Généralisation immédiate à $\sigma_1^2 < \sigma_0^2$ ou à $\sigma_1^2 \neq \sigma_0^2$

Tests sur la variance d'un échantillon gaussien

Soit $(X_1, ..., X_n)$ un échantillon gaussien de loi $\mathcal{N}(m, \sigma^2)$ avec une moyenne m inconnue et les hypothèses

$$H_0 : \sigma^2 = \sigma_0^2 \text{ et } H_1 : \sigma^2 = \sigma_1^2 \text{ avec } \sigma_1^2 > \sigma_0^2 \tag{7}$$

- Définition du test

$$\text{Rejet de } H_0 \text{ si } T = \frac{1}{\sigma_0^2} \sum_{i=1}^{n} (X_i - \overline{X})^2 > S_\alpha \quad \text{avec} \quad \overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

- Remarques
  - La loi de T sous H_0 est une loi du χ^2_{n-1}, ce qui permet de déterminer S_α en fonction de α.
  - Généralisation immédiate à $\sigma_1^2 < \sigma_0^2$ ou à $\sigma_1^2 \neq \sigma_0^2$

Comparaison d'espérances

Soient $(X_1, ..., X_n)$ et $(Y_1, ..., Y_m)$ deux échantillons gaussiens indépendant de lois respectives $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_2, \sigma_2^2)$ avec des variances σ_1^2 et σ_2^2 connues, et les hypothèses

$$H_0 : m_1 = m_2 \text{ et } H_1 : m_1 > m_2 \tag{8}$$

- Définition du test

  Rejet de H_0 si $T = \dfrac{\overline{X} - \overline{Y}}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} > S_\alpha$ avec $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i, \overline{Y} = \dfrac{1}{m}\displaystyle\sum_{j=1}^{m} Y_j$

- Remarques
  - La loi de T sous H_0 est une loi normale $\mathcal{N}(0, 1)$, ce qui permet de déterminer S_α en fonction de α.
  - Généralisation immédiate à $m_1 < m_2$ ou à $m_1 \neq m_2$

Comparaison d'espérances

Soient $(X_1, ..., X_n)$ et $(Y_1, ..., Y_m)$ deux échantillons gaussiens indépendant de lois respectives $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_2, \sigma_2^2)$ avec une même variance $\sigma_1^2 = \sigma_2^2 = \sigma^2$ inconnue, et les hypothèses

$$H_0 : m_1 = m_2 \text{ et } H_1 : m_1 > m_2 \tag{9}$$

- Définition du test

$$\text{Rejet de } H_0 \text{ si } T = \frac{\overline{X} - \overline{Y}}{S_{n,m}(\boldsymbol{x}, \boldsymbol{y})\sqrt{\frac{1}{n} + \frac{1}{m}}} > S_\alpha$$

  avec

$$S_{n,m}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2 + \sum_{j=1}^{m}(Y_j - \overline{Y})^2}{n+m-2}$$

- Remarques
  - La loi de T sous H_0 est une loi de Student à $n+m-2$ ddl.
  - Appelé test de Student ou t-test.
  - Généralisation immédiate à $m_1 < m_2$ ou à $m_1 \neq m_2$

Comparaison d'espérances

Soient $(X_1, ..., X_n)$ et $(Y_1, ..., Y_m)$ deux échantillons gaussiens indépendant de lois respectives $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_2, \sigma_2^2)$ avec des variances σ_1^2 et σ_2^2 inconnues, et les hypothèses

$$H_0 : m_1 = m_2 \text{ et } H_1 : m_1 > m_2 \tag{10}$$

- Définition du test

$$\text{Rejet de } H_0 \text{ si } T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_n^2(\boldsymbol{x})}{n} + \frac{S_m^2(\boldsymbol{y})}{m}}} > S_\alpha$$

  avec

$$S_n^2(\boldsymbol{x}) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \text{ et } S_m^2(\boldsymbol{y}) = \frac{1}{m-1}\sum_{j=1}^{m}(Y_j - \overline{Y})^2$$

- Remarques
  - Sous l'hypothèse H_0, T converge en loi vers une loi normale $\mathcal{N}(0,1)$ lorsque n et m tendent vers ∞.
  - Généralisation immédiate à $m_1 < m_2$ ou à $m_1 \neq m_2$

Comparaison d'espérances

Remarques

- Les tests précédents supposent que les deux échantillons $(X_1, ..., X_n)$ et $(Y_1, ..., Y_m)$ sont indépendants. Si ce n'est pas le cas et que $n = m$, on parle de données appariées. On peut alors considérer les différences $Z_i = X_i - Y_i$ et tester la nullité de l'espérance des Z_i.
- Si l'hypothèse de gaussiannité n'est pas satisfaite, on pourra effectuer un test non paramétrique.

Comparaison de variances

Soient $(X_1, ..., X_n)$ et $(Y_1, ..., Y_m)$ deux échantillons gaussiens indépendant de lois respectives $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_2, \sigma_2^2)$ avec des moyennes m_1 et m_2 connues, et les hypothèses

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ et } H_1 : \sigma_1^2 > \sigma_2^2 \tag{11}$$

- Définition du test

$$\text{Rejet de } H_0 \text{ si } T = \frac{\widetilde{S}_n^2(\boldsymbol{x})}{\widetilde{S}_m^2(\boldsymbol{y})} > S_\alpha$$

avec

$$\widetilde{S}_n^2(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - m_1)^2 \text{ et } \widetilde{S}_m^2(\boldsymbol{y}) = \frac{1}{m}\sum_{j=1}^{m}(Y_j - m_2)^2$$

- Remarques
  - Sous l'hypothèse H_0, T est distribuée suivant une loi de Fisher $\mathcal{F}(n, m)$
  - Appelé F-test.
  - Généralisation immédiate à $\sigma_1^2 < \sigma_2^2$ ou à $\sigma_1^2 \neq \sigma_2^2$

Comparaison de variances

Soient $(X_1, ..., X_n)$ et $(Y_1, ..., Y_m)$ deux échantillons gaussiens indépendant de lois respectives $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_2, \sigma_2^2)$ avec des moyennes m_1 et m_2 inconnues, et les hypothèses

$$H_0 : \sigma_1^2 = \sigma_1^2 \text{ et } H_1 : \sigma_1^2 > \sigma_2^2 \tag{12}$$

- Définition du test

$$\text{Rejet de } H_0 \text{ si } T = \frac{S_n^2(\boldsymbol{x})}{S_m^2(\boldsymbol{y})} > S_\alpha$$

avec

$$S_n^2(\boldsymbol{x}) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \text{ et } S_m^2(\boldsymbol{y}) = \frac{1}{m-1}\sum_{j=1}^{m}(Y_j - \overline{Y})^2$$

- Remarques
  - Sous l'hypothèse H_0, T est distribuée suivant une loi de Fisher $\mathcal{F}(n-1, m-1)$
  - Généralisation immédiate à $\sigma_1^2 < \sigma_2^2$ ou à $\sigma_1^2 \neq \sigma_2^2$

Plan du cours

- Chapitre 1 : Convergence et théorèmes limites
- Chapitre 2 : Estimation
- Chapitre 3 : Tests Statistiques
  - Généralités, exemple
  - Courbes COR
  - Théorème de Neyman Pearson
  - Autres tests paramétriques
  - Test du χ^2
  - Test de Kolmogorov

Test du χ^2

Le test du χ^2 est un test non paramétrique d'ajustement (ou d'adéquation) qui permet de tester les deux hypothèses suivantes

$$H_0 : L = L_0, \quad H_1 : L \neq L_0$$

où L_0 est une loi donnée. Le test consiste à déterminer si $(x_1, ..., x_n)$ est de loi L_0 ou non. On se limitera dans ce cours au cas simple où $x_i \in \mathbb{R}$.

- Définition

$$\text{Rejet de } H_0 \text{ si } \phi_n = \sum_{k=1}^{K} \frac{(Z_k - np_k)^2}{np_k} > S_\alpha$$

- Remarque : L_0 peut être une loi discrète ou continue

Test du χ^2

- Statistique de test
  - Z_k : nombre d'observations x_i appartenant à la classe C_k, $k = 1, ..., K$
  - p_k : probabilité qu'une observation x_i appartienne à la classe C_k sachant $X_i \sim L_0$

$$P\left[X_i \in C_k | X_i \sim L_0\right]$$

  - n : nombre total d'observations
- Loi (asymptotique) de la statistique de test sous H_0

$$\phi_n \underset{n \to \infty}{\overset{\mathcal{L}}{\rightarrow}} \chi^2_{K-1}$$

Remarques

- Interprétation de ϕ_n

$$\phi_n = \sum_{k=1}^{K} \frac{n}{p_k} \left(\frac{Z_k}{n} - p_k \right)^2$$

Distance entre probabilités théoriques et empiriques

- Loi asymptotique de ϕ_n : voir notes de cours ou livres
- Nombre d'observations fini
  Une heuristique dit que la loi asymptotique de ϕ_n est une bonne approximation pour n fini si 80% des classes vérifient $np_k \geq 5$ et si $p_k > 0, \forall k = 1, ..., K$
  ☞ Classes équiprobables

Remarques

- Correction
  Lorsque les paramètres de la loi L_0 sont inconnus

$$\phi_n \underset{n\to\infty}{\overset{\mathcal{L}}{\to}} \chi^2_{K-1-n_p}$$

  où n_p est le nombre de paramètres inconnus estimés par la méthode du maximum de vraisemblance
- Constitution des classes dans le cas d'une loi discrète
- Puissance du test
  Non calculable

Exemple

4.13	1.41	-1.16	-0.75	1.96	2.46	0.197	0.24	0.42	2.00
2.08	1.48	1.73	0.82	0.33	-0.76	0.42	4.60	-2.83	0.197
2.59	0.54	4.06	-0.69	4.99	0.67	2.45	5.61	2.13	1.76
5.03	0.85	1.29	0.17	-0.38	2.76	-1.03	1.87	4.48	0.73

Est-il raisonnable de penser que ces observations sont issues d'une population de loi $\mathcal{N}(1, 4)$?

Solution

- Classes

$$C_1 :]-\infty, -0.34], C_2 :]-0.34, 1], C_3 :]1, 2.34], C_4 :]2.34, \infty[$$

- Nombres d'observations

$$Z_1 = 7,\ Z_2 = 12,\ Z_3 = 10,\ Z_4 = 11$$

Exemple

- Statistique de test

$$\phi_n = 1.4$$

- Seuils

	χ^2_2	χ^2_3
$S_{0.05}$	5.991	7.815
$S_{0.01}$	9.210	11.345

donc on accepte l'hypothèse H_0 avec les risques $\alpha = 0.01$ et $\alpha = 0.05$.

Plan du cours

- Chapitre 1 : Convergence et théorèmes limites
- Chapitre 2 : Estimation
- Chapitre 3 : Tests Statistiques
  - Généralités, exemple
  - Courbes COR
  - Théorème de Neyman Pearson
  - Autres tests paramétriques
  - Test du χ^2
  - Test de Kolmogorov

Test de Kolmogorov

Le test de Kolmogorov est un test non paramétrique d'ajustement (ou d'adéquation) qui permet de tester les deux hypothèses suivantes

$$H_0 : L = L_0, \quad H_1 : L \neq L_0$$

où L_0 est une loi donnée. Le test consiste à déterminer si $(x_1, ..., x_n)$ est de loi L_0 ou non. On se limitera dans ce cours au cas simple où $x_i \in \mathbb{R}$.

- Définition

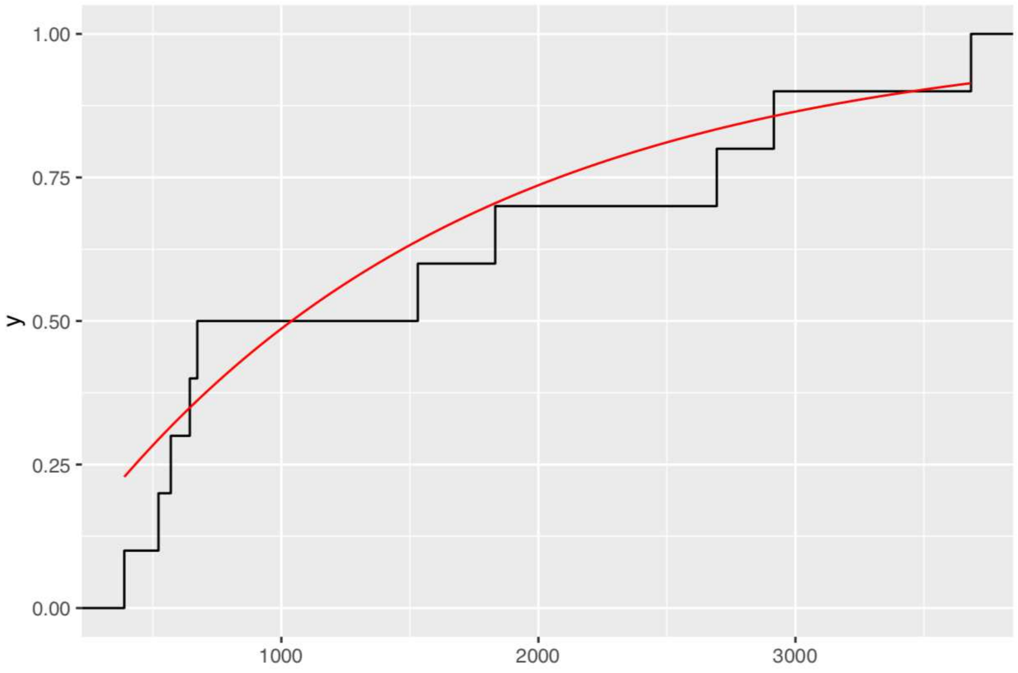
$$\text{Rejet de } H_0 \text{ si } D_n = \sup_{x \in \mathbb{R}} |\widehat{F}(x) - F_0(x)| > S_\alpha$$

- Remarque : L_0 doit être une loi continue

Statistique de test

- Fonctions de répartition

  $F_0(x) = P[X \leq x]$ est la fonction de répartition théorique de L_0 et $\widehat{F}_n(x)$ est la fonction de répartition empirique de $(x_1, ..., x_n)$



D_n est l'écart maximum entre les deux courbes.

Calcul de D_n

- Á l'aide de l'échantillon ordonné

$$D_n = \max_{i \in \{1,...,n\}} \max\{E_i^+, E_i^-\}$$

$$E_i^+ = \left| \widehat{F}_n \left(x_{(i)}^+ \right) - F_0 \left(x_{(i)} \right) \right|, \; E_i^- = \left| \widehat{F}_n \left(x_{(i)}^- \right) - F_0 \left(x_{(i)} \right) \right|$$

  - Rq : $x_{(1)}, ..., x_{(n)}$ est la statistique d'ordre de $x_1, ..., x_n$ telle que $x_{(1)} \leq ... \leq x_{(n)}$
  - Rq : $\widehat{F}_n \left(x_{(i)}^+ \right) = i/n$ et $\widehat{F}_n \left(x_{(i)}^- \right) = (i-1)/n$.

Statistique de test

- Loi de D_n sous H_0
  - Indépendante de L_0
  - Loi asymptotique

$$P[\sqrt{n}D_n < y] \underset{n\to\infty}{\rightarrow} \sum_{k=-\infty}^{+\infty} (-1)^k \exp(-2k^2y^2) = K(y)$$

Convergence de cette série très rapide (pour $y > 0.56$, les trois premiers termes donnent une approximation avec une erreur inférieure à 10^{-4}).

- Détermination du seuil S_α

$$S_{n,\alpha} = \frac{1}{\sqrt{n}} K^{-1}(1-\alpha)$$

Le seuil dépend de α et de n.

Remarques

- Puissance du test

  Non calculable

- Tests unilatéraux

  - Pour tester $H_0 : F = F_0$ contre $H_1 : F \geq F_0$, le test de Kolmogorov rejette H_0 si

    $$D_n^+ = \sup_{t \in \mathbb{R}} \left[\widehat{F}_n(t) - F_0(t) \right] \geq S_{n,\alpha}$$

  - Pour tester $H_0 : F = F_0$ contre $H_1 : F \leq F_0$, le test de Kolmogorov rejette H_0 si

    $$D_n^- = \sup_{t \in \mathbb{R}} \left[F_0(t) - \widehat{F}_n(t) \right] \geq S_{n,\alpha}$$

Exemple

Est-il raisonnable de penser que ces observations sont issues d'une population de loi uniforme sur $[0, 1]$?

x_i	0.0078	0.063	0.10	0.25	0.32	0.39	0.40	0.48	0.49	0.53
E_i^-	0.0078	0.013	0.00	0.10	0.07	0.14	0.05	0.008	0.04	0.03
E_i^+	0.0422	0.037	0.05	0.05	0.12	0.09	0.10	0.13	0.09	0.08
Max(E_i^+, E_i^-)	0.0422	0.037	0.05	0.1	0.12	0.14	0.10	0.13	0.09	0.08

x_i	0.67	0.68	0.69	0.73	0.79	0.80	0.87	0.88	0.90	0.996
E_i^-	0.17	0.13	0.04	0.03	0.04	0.05	0.07	0.03	0.05	0.046
E_i^+	0.12	0.08	0.09	0.08	0.09	0.00	0.02	0.02	0.00	$4e-3$
Max(E_i^+, E_i^-)	0.17	0.13	0.09	0.08	0.09	0.05	0.07	0.03	0.05	0.046

Exemple

- Statistique de test

$$D_n = 0.17$$

- Seuils pour $n = 20$

$S_{0.05}$	0.294
$S_{0.01}$	0.352

donc on accepte l'hypothèse H_0 avec les risques $\alpha = 0.01$ et $\alpha = 0.05$.

Que faut-il savoir ?

- Définition et calcul des risques de première et seconde espèce et de la puissance d'un test binaire
- Définition et détermination des courbes COR
- Appliquer le théorème de Neyman-Pearson dans le cas de variables aléatoires discrètes et continues
- Connaître l'existence des tests paramétriques pour tester la valeur de la moyenne ou de la variance d'un échantillon gaussien
- Connaître l'existence des tests paramétriques pour tester l'égalité de moyennes et de variances pour deux échantillons gaussiens indépendants
- Principe et mise en oeuvre d'un test du χ^2
- Principe et mise en oeuvre d'un test de Kolmogorov