Statistique

Jean-Yves Tourneret[1]

(1) Université of Toulouse, ENSEEIHT-IRIT-TéSA

Thème 1 : Analyse et Synthèse de l'Information

jyt@n7.fr

Plan du cours

- Chapitre 1 : Estimation
  - Modèle statistique, qualités d'un estimateur, exemples
  - Inégalité de Cramér Rao
  - Maximum de vraisemblance
  - Méthode des moments
  - Estimation Bayésienne
  - Intervalles de confiance
- Chapitre 2 : 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.

Modèle Statistique

- Observations

$$x_1, ..., x_n$$

- Échantillon

$$X_1, ..., X_n$$

n va iid associées aux observations

- Estimateur

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

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 $\lim_{n\to+\infty} b_n(\theta) = \lim_{n\to+\infty} v_n(\theta) = 0$

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

Plan du cours

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

Inégalité de Cramér Rao

- Vraisemblance

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

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

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

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

Plan du cours

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

Méthode du Maximum de Vraisemblance

- Définition

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

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

  Si $L(X_1, ..., X_n; \theta)$ est deux fois dérivable et si les bornes de la loi de X_i sont indépendantes de θ

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

  On vérifie qu'on a bien un maximum en étudiant

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

Méthode du Maximum de Vraisemblance

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

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

  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

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

- 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

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

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

Plan du cours

- Chapitre 1 : Estimation
  - Modèle statistique, qualités d'un estimateur, exemples
  - Inégalité de Cramér Rao
  - Maximum de vraisemblance
  - Méthode des moments
  - Estimation Bayésienne
  - Intervalles de confiance
- Chapitre 2 : 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 : Estimation
  - Modèle statistique, qualités d'un estimateur, exemples
  - Inégalité de Cramér Rao
  - Maximum de vraisemblance
  - Méthode des moments
  - Estimation Bayésienne
  - Intervalles de confiance
- Chapitre 2 : 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 $c\left(\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}}\right) = E\left[\left(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\right)^2\right]$ 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

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

  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

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

Plan du cours

- Chapitre 1 : Estimation
  - Modèle statistique, qualités d'un estimateur, exemples
  - Inégalité de Cramér Rao
  - Maximum de vraisemblance
  - Méthode des moments
  - Estimation Bayésienne
  - Intervalles de confiance
- Chapitre 2 : 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 3 : $X_i \sim \mathcal{N}(m, \sigma^2)$, 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.

Plan du cours

- Chapitre 1 : Estimation
- Chapitre 2 : Tests Statistiques
  - Généralités, exemple
  - Courbes COR
  - Théorème de Neyman Pearson
  - Test du rapport de vraisemblance généralisé
  - Test du χ^2
  - Test de Kolmogorov

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

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

- 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 : Estimation
- Chapitre 2 : Tests Statistiques
  - Généralités, exemple
  - Courbes COR
  - Théorème de Neyman Pearson
  - Test du rapport de vraisemblance généralisé
  - 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

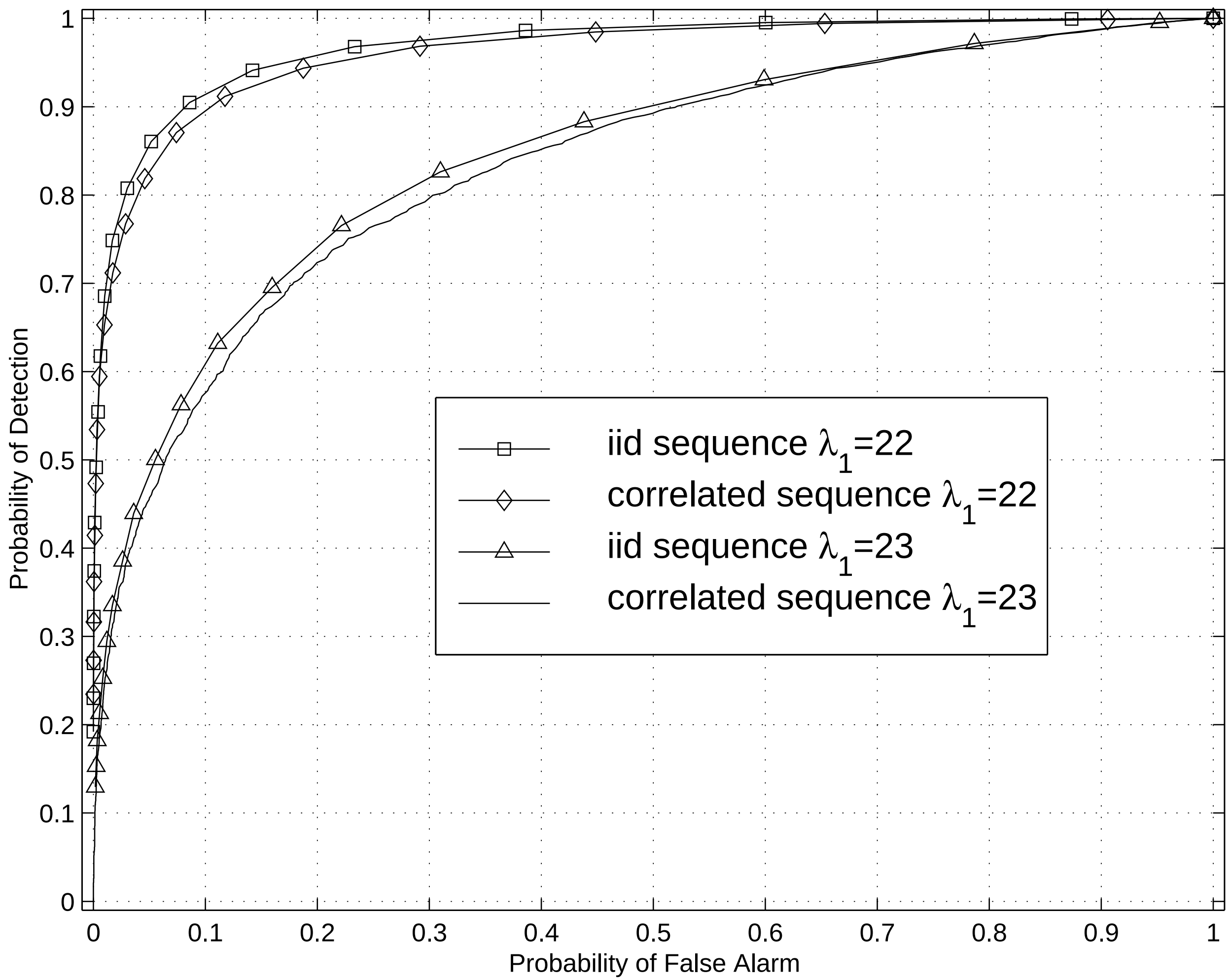
$$S_\alpha = m_0 + \frac{\sigma}{\sqrt{n}} F^{-1}(1-\alpha)$$

  - Probabilité de détection

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

Représentation graphique

ROC's for iid and correlated sequences



Plan du cours

- Chapitre 1 : Estimation
- Chapitre 2 : Tests Statistiques
  - Généralités, exemple
  - Courbes COR
  - Théorème de Neyman Pearson
  - Test du rapport de vraisemblance généralisé
  - Test du χ^2
  - Test de Kolmogorov

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 : Estimation
- Chapitre 2 : Tests Statistiques
  - Généralités, exemple
  - Courbes COR
  - Théorème de Neyman Pearson
  - Test du rapport de vraisemblance généralisé
  - 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

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

Plan du cours

- Chapitre 1 : Estimation
- Chapitre 2 : Tests Statistiques
  - Généralités, exemple
  - Courbes COR
  - Théorème de Neyman Pearson
  - Test du rapport de vraisemblance généralisé
  - 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 de la statistique de test

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

Remarques

- Interprétation de ϕ

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

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

$$\phi_n \xrightarrow[n \to \infty]{\mathcal{L}} \chi^2_{K-1-n_p}$$

  où n_p est le nombre de paramètres inconnus estimés par la méthode du maximum de vraisemblance
- 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 : Estimation
- Chapitre 2 : Tests Statistiques
  - Généralités, exemple
  - Courbes COR
  - Théorème de Neyman Pearson
  - Test du rapport de vraisemblance généralisé
  - 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

Remarques

- Statistique de test
  $F_0(x)$ est la fonction de répartition théorique associée à L_0 et $\widehat{F}(x)$ est la fonction de répartition empirique de $(x_1, ..., x_n)$
- Loi asymptotique de D_n : voir livres

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

- Détermination du seuil S_α : $S_\alpha = \frac{1}{\sqrt{n}} K^{-1}(1-\alpha)$
  Le seuil dépend de α et de n.

Remarques

- Calcul de D_n

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

$$E_i^+ = \left| \widehat{F}\left(x_i^{*+}\right) - F_0\left(x_i^*\right) \right|, \; E_i^- = \left| \widehat{F}\left(x_i^{*-}\right) - F_0\left(x_i^*\right) \right|$$

  - Rq : $x_1^*, ..., x_n^*$ est la statistique d'ordre de $x_1, ..., x_n$.
  - Rq : $\widehat{F}\left(x_i^{*+}\right) = i/n$ et $\widehat{F}\left(x_i^{*-}\right) = (i-1)/n$.
- Puissance du test
  Non calculable

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 $\alpha = 0.05$.