Statistiques pour l'Ingénieur

Cours n°5/6

Statistique inférentielle : Estimation

- définitions
- qualités d'un estimateur
- théorème de Cramer-Rao

Ph. LERAY - A. ROGOZAN

Estimation : formalisation et notations

- Soit une v.a. X , dont la loi - supposée connue - dépend d'un paramètre Θ qui est inconnu
  - NB: Θ peut être vectoriel
- Supposons que l'on dispose d'une réalisation $x_1,\ldots,x_n$ d'un échantillon i.i.d. $X_1,\ldots,X_n$ de loi parente X

- Objectif : Estimer Θ à partir de la réalisation $x_1,\ldots,x_n$

- L'estimation peut être :
  - ponctuelle : Θ = valeur
  - par intervalles : $\Theta \in [a,b]$

Estimation : formalisation et notations

- Objectif : Estimer Θ à partir de la réalisation $x_1,\ldots,x_n$

Théorique

- Estimateur du paramètre Θ :
  - $T = f(X_1,\ldots,X_n)$ que l'on note aussi $\hat{\Theta}$
  - T = statistique de l'échantillon
    = v.a. ayant sa propre loi de probabilité

Empirique

- Estimation :
  - réalisation $t=f(x_1,\ldots,x_n)$ de l'estimateur $T=f(X_1,\ldots,X_n)$

- Questions :
  - Comment construire la statistique T ?
  - Quelle est la précision de l'estimation t ?
  - Comment choisir la taille de l'échantillon n ?

Qualités d'un estimateur

- Problème : un même paramètre Θ peut être estimé à l'aide d'estimateurs différents
  - ex : pour une distribution symétrique, la médiane et la moyenne empirique sont des estimations de l'espérance
- Il faut définir les qualités exigées d'un estimateur afin de choisir entre plusieurs estimateurs possibles
  - estimateur convergent : quand $T \xrightarrow[n \to \infty]{P} \Theta$ c.a.d. si
    $$\forall \epsilon > 0, \lim_{n \to \infty} P(|T - \Theta| > \epsilon) = 0$$
  - ex : la loi de grands nombres permet de montrer que $\bar{X}$ est un estimateur convergent de μ

Qualités d'un estimateur

- Problème : deux estimateurs convergents ne convergent pas à la même vitesse vers le paramètre Θ
  - la vitesse de convergence dépend de *n*
  - à *n* fixé, elle dépend de la précision de l'estimateur

- La précision (ou risque) de l'estimateur dépend de l'erreur d'estimation :
  - Erreur = $T - \Theta$ (v.a.)
  - Précision = Erreur quadratique moyenne (EQM) :

$$\boxed{R(T,\Theta) = E\left[(T-\Theta)^2\right] = Var(T) + (E(T) - \Theta)^2}$$

= var. de l'estimateur + biais^2

Qualités d'un estimateur

- Estimateur préférable :

  T_1 préférable à T_2 si $\forall \Theta \in \Re$, $R(T_1, \Theta) \leqslant R(T_2, \Theta)$

- Estimateur admissible :

  T admissible s'il n'existe pas d'autre estimateur qui lui soit préférable

Biais d'un estimateur : $E(T)-\Theta$

- Stratégie MIN-MAX : $\min_T \{ \max_\Theta \{ R(T,\Theta) \} \}$
  - Minimiser l'EQM = minimiser $E[(T-\Theta)^2]$, souvent difficile
- Approche : chercher parmi les estimateurs sans biais
  - Trouver l'estimateur sans biais de variance minimale

- Paramètre : $\mu = E(X)$
- Estimateur : $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
  - SANS BIAIS car $E(\bar{X}) = E(X) = \mu$

- Paramètre : $\sigma^2 = Var(X)$
- Estimateur : $S^2 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2$
  - AVEC BIAIS car $E(S^2) = \frac{(n-1)}{n} Var(X) \neq \sigma^2$
  - Estimateur sans biais : $S^{*2} = \frac{n}{n-1} S^2$

Exemple d'estimateur biaisé S^2

- Paramètre : $\sigma^2 = Var(X)$ estimateur : $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$

- Montrez le th. de Huygens

$$\forall a \in \mathbb{R}, \sum_{i=1}^{n}(X_i - a)^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2 + n(\bar{X} - a)^2$$

- Utilisez-le avec *a*=*μ* et montrez que $E(S^2) = \frac{(n-1)}{n} Var(X) \neq \sigma^2$

- Montrez que $S^{*2} = \frac{n}{n-1} S^2$ n'est pas biaisé

Qualités d'un estimateur

- Propriété :

Si T est un estimateur (asymptotiquement) sans biais et si

$\lim_{n \to \infty} E(T) = \Theta$ et $\lim_{n \to \infty} Var(T) = 0$ alors T est convergent.

  - Remarque : la convergence et l'absence de biais ne garantissent pas l'unicité d'un estimateur

- Problème : soit une v.a. X
  - définie par une famille paramétrée de lois $f(x;\Theta)$ [f connue]
  - observée sur $x_1,\dots,x_n$
  - comment trouver l'estimateur T sans biais de Θ de variance minimum ?

indice = notion de statistique exhaustive

Statistique exhaustive

- Soient
  - $X_1,\ldots,X_n$ un n-échantillon d'une v.a.
  - $L(x_1,,x_n,\Theta)$ la vraisemblance de l'échantillon
  - T une statistique

- Statistique exhaustive : T est dite exhaustive
  - si elle contient toute « l'information » apportée par l'échantillon relativement à la connaissance de Θ
  - i.e. par le principe de factorisation, si ∃ g et h /

$$L(x_1,,x_n,\Theta)=g(t,\Theta)h(x_1,,x_n)$$

  - remarque : la factorisation permet de reconnaître que T (connue) est exhaustive, mais permet difficilement de la construire ou même de savoir s'il en existe une

Exemples de statistiques exhaustives

- Loi de Bernoulli de paramètre inconnu *p*
  - $T=\sum_{i=1}^{n} X_i$ est exhaustive pour *p*

$L(x_1,\ldots,x_n,p) = p^t(1-p)^{n-t}$ [h(...)=1]

- Loi exponentielle $\frac{1}{\Theta}\exp\left(\frac{-x}{\Theta}\right)$ de paramètre inconnu Θ
  - $T=\sum_{i=1}^{n} X_i$ est exhaustive pour Θ

$L(x_1,\ldots,x_n,\Theta) = \ldots$ [h(...)=1]

Exemples de statistiques exhaustives

- Loi normale de paramètres μ et σ^2 :
  - Si σ^2 est connu,

    $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ est exhaustive pour le paramètre inconnu μ

    $L(x_1,\ldots,x_n,\mu) = \ldots$

  - Si μ est connu, $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$ est exhaustive pour σ^2

Statistique exhaustive

- Quelles sont les lois permettant une statistique exhaustive ?

Théorème de Darmois

- soit une v.a. X / domaine de définition ne dépend pas de Θ
- soit $X_1,\ldots,X_n$ un échantillon
- une condition nécessaire et suffisante pour que l'échantillon admette une statistique suffisante est que la densité de X appartienne à la famille exponentielle :

$$f(x,\Theta) = \exp\left[\, a(x)\alpha(\Theta) + b(x) + \beta(\Theta) \,\right]$$

- de plus, si a bijective et continûment différentiable alors

$$T = \sum_{i=1}^{n} a(X_i)$$ est une statistique exhaustive

Statistique exhaustive

- Exceptions :
  - il est possible de trouver des statistiques exhaustives pour la loi uniforme même si le théorème de Darmois ne s'applique pas

Information de Fisher

- Définition : $I_n(\Theta)=E[(\frac{\partial \ln L(x_1,,x_n,\Theta)}{\partial \Theta})^2]$

  quantité d'information de Fisher apportée par un échantillon sur le paramètre Θ [si elle existe]

- Propriétés (si le domaine de définition de X ne dépend pas de Θ)
  - $I_n(\Theta)=-E(\frac{\partial^2 \ln L}{\partial \Theta^2})$
  - additivité : $I_n(\Theta)=n I_1(\Theta)$ chaque observation apporte autant d'info.
  - dégradation de l'information : $I_T(\Theta)\leq I_n(\Theta)$

    une statistique T apporte une quantité d'info inférieure ou égale à celle apportée par l'échantillon ... il y a égalité si T est exhaustive

Information de Fisher

- Exemple :
  - Calculer l'information apportée par une observation sur le paramètre μ d'une loi normale ou σ^2 est connu
  - On retrouve le fait que chaque observation apporte d'autant plus d'information que la dispersion est petite

Retour à nos estimateurs ...

- Problème : soit une v.a. X
  - définie par une famille paramétrée de lois $f(x,\Theta)$ [f connue]
  - observée sur $x_1,\ldots,x_n$
  - comment trouver l'estimateur T sans biais de Θ de variance minimale ?

indice = notion de statistique exhaustive

- Une réponse : le théorème de Cramer-Rao
  - qui sous certaines conditions nous permet de
    - calculer la borne inférieure de la variance d'un estimateur sans biais
    - construire un estimateur efficace (i.e. celui qui atteindra cette borne min.)

Estimateur efficace

- Estimateur efficace = estimateur sans biais
  - vérifiant l'ensemble des conditions de Cramer-Rao
  - dont la variance est égale à la borne de Cramer-Rao
- Propriété :
  - Efficace => sans biais de variance minimale
- Attention :
  - Un estimateur efficace n'EXISTE pas toujours!!!
    (à cause des conditions de Cramer-Rao)
  - Sans biais de variance minimale ~~=>~~ efficace

Théorème de Cramer-Rao

- Soit û un estimateur sans biais de u(Θ), si

conditions de Cramer-Rao

- le support de la loi de X est indépendant de Θ
- $\frac{\partial L}{\partial \Theta}$ existe et est continue par rapport à Θ
- $I_n(\Theta)$ l'information de Fisher est finie
- $\frac{\partial L}{\partial \Theta}$ et $\hat{u}\frac{\partial L}{\partial \Theta}$ intégrables par rapport à Θ

Alors $\boxed{Var(\hat{u}) \geq \frac{(u'(\Theta))^2}{I_n(\Theta)}}$

borne de Cramer-Rao

Estimateur efficace (le retour...)

- Si les conditions de Cramer-Rao sont vérifiées,
  alors û est un estimateur efficace de u(Θ)
  ssi ∃ une fonction A indépendante de $X_1,\ldots,X_n$, telle que :

$$\boxed{\frac{\partial \ln L}{\partial \Theta}(X_1,,X_n,\Theta)=A(n,\Theta)(\hat{u}-u(\Theta))}$$

- De plus, la variance de cet estimateur est : $Var(\hat{u})=\dfrac{u'(\Theta)}{A(n,\Theta)}$

Exemples

- Soit une v.a. X de densité $f_X(x) = \Theta e^{-\Theta x}$ *où* $x \in (0, \infty)$ de paramètre inconnu Θ
  - Il n'existe pas d'*estimateur efficace* de Θ !
  - $\bar{X}$ est un *estimateur efficace* de 1/Θ

Exemples

- Soit une v.a. $X \sim N(\mu, \sigma^2)$, avec μ inconnu et σ^2 connu
  - $\bar{X}$ est un *estimateur efficace* de μ