CONVERGENCE ET ESTIMATION

1) Introduction : Le problème central de l'estimation en statistique est le suivant : disposant d'observations sur un échantillon de taille n on souhaite en déduire les propriétés de la population dont il est issu.
   On cherchera à estimer, par exemple, la moyenne d'une population à partir de la moyenne d'un échantillon. Le mode de tirage le plus important est l'échantillonnage aléatoire simple correspondant à des tirages équiprobables et indépendants les uns des autres.
   L'une des premières qualités d'un estimateur est d'être convergent en probabilité vers le paramètre à estimer. Un échantillon de X est une suite de variables aléatoires $(X_1, X_2, ..., X_n)$ indépendantes et de même loi que X. Un estimateur d'un paramètre θ inconnu est une fonction $T = f(X_1, X_2, ..., X_n)$ qui dépend de l'échantillon et donc T doit converger en probabilité vers le paramètre θ. La précision d'un estimateur sera mesuré par sa variance.

2) Rappels de la loi Gamma et la loi Normale
   On dit qu'une variable aléatoire positive X suit une loi gamma de paramètre r, notée γ_r si sa densité est donnée par : $f(x) = \dfrac{1}{\Gamma(r)} \exp(-x) x^{r-1}$

   Avec $\Gamma(x) = \int_0^{+\infty} \exp(-t) t^{x-1} dt$ (fonction Gamma) définie pour $x > 0$

   Propriétés de la fonction Gamma

   (1) $\Gamma(x+1) = x\Gamma(x)$ (intégration par partie)
   (2) $\Gamma(1) = 1$
   (3) $\Gamma(n+1) = n!$
   (4) $\Gamma(k + \frac{1}{2}) = \dfrac{1.3.5......(2k-1)}{2^k} \Gamma(\frac{1}{2})$
   (5) $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

   Espérance de la loi γ_r : Soit X une variable aléatoire suivant la loi gamma de paramètre r

On a $E(X)=\frac{1}{\Gamma(r)}\int_0^{+\infty} t\exp(-t)t^{r-1}dt=\frac{1}{\Gamma(r)}\int_0^{+\infty} t^r\exp(-t)dt=\frac{\Gamma(r+1)}{\Gamma(r)}=r$

Variance de la loi γ_r : $V(X)=E(X^2)-E^2(X)$

$$E(X^2)=\frac{1}{\Gamma(r)}\int_0^{+\infty} t^2\exp(-t)t^{r-1}dt=\frac{1}{\Gamma(r)}\int_0^{+\infty} t^{r+1}\exp(-t)dt=\frac{\Gamma(r+2)}{\Gamma(r)}=r(r+1)$$

Donc $V(X)=r(r+1)-r^2=r$

Loi Normale de paramètres (m,σ)

On dit qu'une variable aléatoire X suit la loi normale notée $N(m,\sigma)$, si sa densité est $f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{1}{2}(\frac{x-m}{\sigma})^2)$ où $m=E(X)$ et $\sigma=\sqrt{V(X)}$ (écart type)

Avec le changement de variable $U=\frac{X-m}{\sigma}$ (variable normale centrée réduite), la densité de U est $f(u)=\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}u^2)$

Montrons que $V(U)=1$

On a $V(U)=E(U^2)=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}u^2\exp(-\frac{1}{2}u^2)du=\frac{2}{\sqrt{2\pi}}\int_0^{+\infty}u^2\exp(-\frac{1}{2}u^2)du$

Posons $t=\frac{u^2}{2}$, $dt=udu$

$$V(U)=\frac{2}{\sqrt{2\pi}}\int_0^{+\infty}2t\exp(-t)\frac{dt}{\sqrt{2t}}=\frac{2}{\sqrt{\pi}}\int_0^{+\infty}t^{\frac{1}{2}}\exp(-t)dt=\frac{2}{\sqrt{\pi}}\Gamma(\frac{3}{2})=\frac{2}{\sqrt{\pi}}\frac{1}{2}\Gamma(\frac{1}{2})$$

Donc $V(U)=\frac{1}{\sqrt{\pi}}\sqrt{\pi}=1$

Moments de la loi normale centrée réduite

Soit U une variable normale centrée réduite, on appelle moment d'ordre k de U : $\mu_k=E(U^k)$

_Si $k=2p+1$ alors $\mu_{2p+1}=0$ (car fonction impaire)

Si $k=2p$ alors $\mu_{2p}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}u^{2p}\exp(-\frac{1}{2}u^2)du=\frac{2}{\sqrt{2\pi}}\int_0^{+\infty}u^{2p}\exp(\frac{-1}{2}u^2)du$

Posons $t=\frac{u^2}{2}$, $dt=udu$

$$\mu_{2p} = \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} (2t)^p \exp(-t) \frac{dt}{\sqrt{2t}} = \frac{2^p}{\sqrt{\pi}} \int_0^{+\infty} t^{p-\frac{1}{2}} \exp(-t) dt = \frac{2^p}{\sqrt{\pi}} \Gamma(p+\frac{1}{2})$$

Or $\Gamma(p+\frac{1}{2}) = \frac{1.3.5.....(2p-1)}{2^p} \Gamma(\frac{1}{2})$ et $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Donc $\mu_{2p} = 1.3.5......(2p-1) = \frac{(2p)!}{2^p p!}$

3) Fonctions caractéristiques

Définition : la fonction caractéristique d'une variable aléatoire réelle X est la transformée de Fourier de sa loi de probabilité. elle est notée $\varphi_X(t)$ et
On a $\varphi_X(t) = E(\exp(itX))$ (i complexe)

Si X est une variable à densité (X est une v.a continue de densité f) alors :

$$\varphi_X(t) = \int_{IR} \exp(itx) f(x) dx$$

Si X est une variable discrète alors sa fonction caractéristique est :

$$\varphi_X(t) = \sum_k \exp(itk) P(X = k)$$

Propriétés

1) $\varphi_{\lambda X}(t) = \varphi_X(\lambda t)$ $\forall \lambda$ un scalaire
2) $\varphi_{X+a}(t) = \exp(ita)\varphi_X(t)$ $\forall a$ un scalaire
3) Si X est une variable aléatoire d'espérance m et d'écart type σ
   Et $U = \frac{X-m}{\sigma}$

   $$\varphi_{\frac{X-m}{\sigma}}(t) = \varphi_U(t) = \exp(-\frac{itm}{\sigma})\varphi_X(\frac{t}{\sigma})$$

Remarque :la fonction caractéristique se prête bien aux additions de

Variables aléatoires indépendantes :

Si X et Y sont deux variables aléatoires indépendantes alors

$$\varphi_{X+Y}(t) = \varphi_X(t).\varphi_Y(t)$$

En effet $\varphi_{X+Y}(t) = E(\exp(it(X+Y))) = E(\exp(itX)\exp(itY))$

Or X et Y sont indépendantes $E(\exp(itX)\exp(itY)) = E(\exp(itX))E(\exp(ity))$

Donc $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$

<u>Proposition :</u> Soit X une variable aléatoire de fonction de répartition $\varphi_X(t)$

On a $\varphi_X(0) = 1$ et $\dfrac{d^k\varphi_X}{dt^k}(0) = \varphi_X{}^{(k)}(0) = t^k E(X^k)$

Démo : supposons que X est une variable continue de densité f

On a $\varphi_X(t) = \int_{IR} \exp(itx) f(x)dx \Rightarrow \varphi_X(0) = \int_{IR} f(x)dx = 1$ (car f est une densité)

En dérivant $\varphi_X(t)$ par rapport à t : $\varphi'_X(t) = i\int_{IR} x\exp(itx) f(x)dx$

Si t=0 $\varphi'_X(0) = i\int_{IR} xf(x)dx = iE(X)$

Si on dérive 2 fois , $\varphi_X{}^{(2)}(t) = \int_{IR} (ix)^2 \exp(itx) f(x)dx$

Pour t=0, on a $\varphi_X{}^{(2)}(0) = (i)^2 \int_{IR} x^2 f(x)dx = -\int_{IR} x^2 f(x)dx = -E(X^2)$

En dérivant k fois par rapport à t : $\varphi_X{}^{(k)}(t) = \int_{IR} (ix)^k \exp(itx) f(x)dx$

Donc $\varphi_X{}^{(k)}(0) = (i^k)\int_{IR} x^k f(x)dx = i^k E(X^k) \quad \forall k \in IN$

<u>Formule de Mac-Laurin</u>

Si $\varphi_X(t)$ est indéfiniment dérivable on a :

$$\varphi_X(t) = \sum_{k=0}^{+\infty} \frac{t^k}{k!}\varphi_X^{(k)}(0) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} i^k E(X^k)$$

Exemple1 : Soit X une variable aléatoire continue de densité :

$f(x) = \exp(-x)$ si $x > 0$ et $f(x) = 0$ sinon

Déterminer la fonction caractéristique de X

On a $\varphi_X(t) = \int_{IR} \exp(itx) f(x)dx = \int_0^{+\infty} \exp(itx)\exp(-x)dx = \int_0^{+\infty} \exp(-(1-it)x)dx$

$$\varphi_X(t) = \int_0^{+\infty} \exp(-(1-it)x)dx = \left[\frac{-\exp(-(1-it)x)}{(1-it)}\right]_0^{+\infty} = \frac{1}{1-it}$$

Car $\exp(-(1-it)x) = \exp(-x)\exp(itx) \to 0$ lorsque $x \to +\infty$

Puisque $\exp(itx)$ est bornée de module 1 et $\exp(-x) \to 0$ quand $x \to +\infty$

Exemple2 : Déterminer la fonction caractéristique de la loi de Bernoulli de paramètre p

Soit X une variable de Bernoulli :

$X = 1$ avec la probabilité p et $X = 0$ avec la probabilité 1-p

X étant discrète, donc sa fonction caractéristique est :

$$\varphi_X(t) = \sum_k \exp(itk)P(X = k) = \sum_{k=0}^{1} \exp(itk)P(X = k) = P(X = 0) + \exp(it)P(X = 1)$$

$$\varphi_X(t) = 1 - p + p\exp(it) = q + p\exp(it) \text{ avec } q = 1 - p$$

4) <u>Convergences des suites de variables aléatoires</u>

Une suite (X_n) de variables aléatoires étant une suite de fonctions il existe diverses façons de définir la convergence de (X_n) dont certaines jouent un grand rôle en statistiques.

a) <u>Convergence en probabilité</u>

<u>Définition</u>

La suite (X_n) converge en probabilité vers une variable aléatoire X Si $\forall \varepsilon > 0, \eta > 0$ (arbitrairement petits) il existe un entier n_0 tel que $\forall n > n_0 \Rightarrow P(|X_n - X| > \varepsilon) < \eta$

C'est-à-dire $P(|X_n - X| > \varepsilon) \xrightarrow[n \to +\infty]{} 0$

On notera $(X_n) \xrightarrow{P} X$

<u>Inégalité de Bienaymé-Tchebych</u>ev

$$P(|X - E(X)| > \varepsilon) < \frac{V(X)}{\varepsilon^2} \quad \forall \varepsilon > 0$$

Remarque : Lorsque $E(X_n) \xrightarrow[n \to +\infty]{} a$, il suffit de montrer que $V(X_n) \xrightarrow[n \to +\infty]{} 0$ pour établir la convergence en probabilité de La suite (X_n) vers a.

En effet d'après Tchebychev : $P(|X_n - E(X_n)| > \varepsilon) < \frac{V(X_n)}{\varepsilon^2} \to 0$

Donc en passant à la limite : $\lim_{n\to+\infty} P(|X_n - a| > \varepsilon) = 0 \ \forall \varepsilon > 0$

b) <u>Convergence en moyenne quadratique</u>

On suppose que $E(|X_n - X|^2)$ existe

<u>Définition</u>

On dit qu'une suite de variables aléatoires (X_n) converge en moyenne quadratique vers une variable X si $E(|X_n - X|^2) \xrightarrow[n\to+\infty]{} 0$

on notera $(X_n) \xrightarrow{m.q} X$

c) <u>Convergence en loi</u>

<u>Définition</u>

La suite (X_n) converge en loi vers la variable X de fonction de répartition F si en tout point de continuité de F la suite (F_n) des fonctions de répartition des (X_n) converge vers F

C'est-à-dire $\lim_{n\to+\infty} F_n(x) = F(x)$ pour tout x point de continuité de F

On notera $(X_n) \xrightarrow{L} X$

<u>Remarque :</u> Pour les variables discrètes, la convergence en loi est équivalente à $\lim_{n\to+\infty} P(X_n = k) = P(X = k)$

<u>Théorème :</u>Si la suite des fonctions caractéristiques $\varphi_{X_n}(t)$ converge vers $\varphi_X(t)$ alors $(X_n) \xrightarrow{L} X$

5) <u>Applications</u>

<u>Convergence en loi de la binomiale vers la loi Normale</u>

<u>Théorème (Moivre-laplace)</u>

Soit (X_n) une suite de variables binomiales $B(n,p)$

Alors $\frac{X_n - np}{\sqrt{npq}} \xrightarrow{L} N(0,1)$ lorsque $n \to +\infty$

Démonstration : la fonction caractéristique de la loi $B(n,p)$ est :

$\varphi_{X_n}(t) = (p\exp(it) + 1 - p)^n$ donc celle de $Y_n = \dfrac{X_n - np}{\sqrt{npq}}$ est :

$$\varphi_{Y_n}(t) = (p\exp(\frac{it}{\sqrt{npq}}) + 1 - p)^n \exp(\frac{-itnp}{\sqrt{npq}})$$

$$Ln(\varphi_{Y_n}(t)) = nLn(p(\exp(\frac{it}{\sqrt{npq}}) - 1) + 1) - \frac{itnp}{\sqrt{npq}}$$

On rappelle le développement limité de l'exponentielle à l'ordre 2

$\exp(x) \approx 1 + x + \dfrac{x^2}{2}$ (au voisinage de 0)

$$Ln(\varphi_{Y_n}(t)) \approx nLn(p(\frac{it}{\sqrt{npq}} - \frac{t^2}{2npq}) + 1) - \frac{itnp}{\sqrt{npq}}$$

On rappelle $Ln(1+x) \approx x - \dfrac{x^2}{2}$ (au voisinage de 0)

Donc $Ln(\varphi_{Y_n}(t)) \approx n[\dfrac{pit}{\sqrt{npq}} - \dfrac{pt^2}{2npq} + \dfrac{p^2t^2}{2npq}] - \dfrac{itnp}{\sqrt{npq}}$

$$Ln(\varphi_{Y_n}(t)) \approx -\frac{t^2}{2q} + \frac{pt^2}{2q} = \frac{t^2}{2q}(p-1) = -\frac{t^2}{2}$$

En composant par l'exponentielle :

$\varphi_{Y_n}(t) \approx \exp(-\dfrac{t^2}{2})$ fonction caractéristique de la loi normale N(0,1)

Conclusion : $\dfrac{X_n - np}{\sqrt{npq}} \xrightarrow{L} N(0,1)$

<u>Remarque :</u>lorsque n est assez grand on peut donc approximer la loi Binomiale par la loi normale. On donne généralement comme Condition np et $nq > 5$

Il convient cependant d'effectuer la correction de continuité :

On obtient donc une valeur approchée de P(X=x) par la surface sous La courbe de densité de la loi normale $N(np, \sqrt{npq})$ comprise entre Les droites d'abscisse $x - \dfrac{1}{2}$ et $x + \dfrac{1}{2}$

$$P(X = x) \approx P(x - \frac{1}{2} < X < x + \frac{1}{2}) = P(\frac{x - \frac{1}{2} - np}{\sqrt{npq}} < \frac{X - np}{\sqrt{npq}} < \frac{x + \frac{1}{2} - np}{\sqrt{npq}})$$

Et $$P(X \le x) \approx P(\frac{X-np}{\sqrt{npq}} < \frac{x+\frac{1}{2}-np}{\sqrt{npq}})$$

Exemple : Soit X une variable binomiale B(n=40 ;p=0,3)

La valeur exacte pour $P(X=11)$ est 0,1319

La formule d'approximation :

$$P(X=11) \approx P(\frac{11-\frac{1}{2}-12}{\sqrt{8{,}4}} < \frac{X-12}{\sqrt{8{,}4}} < \frac{11+\frac{1}{2}-12}{\sqrt{8{,}4}}) = P(-0{,}52 < U < -0{,}17) = 0{,}131$$

Avec $np = 12$ et $npq = 8{,}4$

Donc l'erreur est de moins de 1%

Convergence en loi de la loi de Poisson vers la loi normale

Théorème : Soit (X_λ) une suite de variables de Poisson de paramètre λ

Alors si $\lambda \to +\infty$, $\frac{X_\lambda - \lambda}{\sqrt{\lambda}} \xrightarrow{L} N(0,1)$

Démo : on rappelle la fonction caractéristique de la loi de Poisson :

$$\varphi_{X_\lambda}(t) = \exp(\lambda \exp(it) - \lambda)$$

On rappelle aussi la formule $\varphi_{\frac{X-m}{\sigma}}(t) = \exp(-\frac{itm}{\sigma})\varphi_X(\frac{t}{\sigma})$

Donc la fonction caractéristique de la variable $\frac{X_\lambda - \lambda}{\sqrt{\lambda}}$ est :

$$\varphi_{\frac{X_\lambda-\lambda}{\sqrt{\lambda}}}(t) = \exp(-\frac{it\lambda}{\sqrt{\lambda}})\varphi_X(\frac{t}{\sqrt{\lambda}}) = \exp(-\frac{it\lambda}{\sqrt{\lambda}})\exp(\lambda\exp(i\frac{t}{\sqrt{\lambda}}) - \lambda)$$

En utilisant le développement limité à l'ordre 2 de la fonction expo

$$\exp(\frac{it}{\sqrt{\lambda}}) \approx 1 + \frac{it}{\sqrt{\lambda}} - \frac{t^2}{2\lambda}$$

Donc $\varphi_{\frac{X_\lambda-\lambda}{\sqrt{\lambda}}}(t) = \exp(-\frac{it\lambda}{\sqrt{\lambda}} + \lambda + \frac{\lambda it}{\sqrt{\lambda}} - \frac{t^2}{2} - \lambda) = \exp(-\frac{t^2}{2})$

On retrouve la fonction caractéristique de la loi normale centrée et réduite.

Conclusion : $\frac{X_\lambda - \lambda}{\sqrt{\lambda}} \xrightarrow{L} N(0,1)$.

Théorème (Central-limite)

Soit (X_n) une suite de variables aléatoires, indépendantes et de même Loi d'espérance m et d'écart-type σ alors :

$$\frac{X_1 + X_2 + \ldots\ldots + X_n - nm}{\sigma\sqrt{n}} \xrightarrow{L} N(0,1)$$

Démonstration : $\dfrac{X_1 + X_2 + \ldots\ldots + X_n - nm}{\sigma\sqrt{n}} = \displaystyle\sum_{i=1}^{n} \frac{X_i - m}{\sigma\sqrt{n}}$

Posons $Y_n = \displaystyle\sum_{i=1}^{n} \frac{X_i - m}{\sigma\sqrt{n}}$

$$E\left(\frac{X_i - m}{\sigma\sqrt{n}}\right) = \frac{E(X_i) - m}{\sigma\sqrt{n}} = 0 \quad \text{et} \quad V\left(\frac{X_i - m}{\sigma\sqrt{n}}\right) = \frac{1}{\sigma^2 n} V(X_i) = \frac{\sigma^2}{n\sigma^2} = \frac{1}{n}$$

La fonction caractéristique de $Y_n = \displaystyle\sum_{i=1}^{n} \frac{X_i - m}{\sigma\sqrt{n}}$ est :

$$\varphi_{Y_n}(t) = \prod_{i=1}^{n} \varphi_{\frac{X_i - m}{\sigma\sqrt{n}}}(t) = \left(\varphi_{\frac{X_i - m}{\sigma\sqrt{n}}}(t)\right)^n = \left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n^2}\right)\right)^n$$

On rappelle que $\left(1 + \dfrac{x}{n}\right)^n \to \exp(x)$ lorsque $n \to +\infty$

Car $\left(1 + \dfrac{x}{n}\right)^n = e^{n\ln(1+\frac{x}{n})} \approx e^{n\frac{x}{n}} = e^x$

Donc $\varphi_{Y_n}(t) = \left(1 - \dfrac{t^2}{2n} + o\left(\dfrac{1}{n^2}\right)\right)^n \to e^{-\frac{t^2}{2}}$ lorsque $n \to +\infty$

Ce qui montre la convergence en loi vers la loi normale

6) Estimateurs

Définition : Soit $(X_1, X_2, \ldots\ldots, X_n)$ un échantillon de X c'est-à-dire une suite de variables aléatoires indépendantes et de même loi que X

La statistique $\overline{X}$ ou moyenne empirique de l'échantillon est $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$

$$E(\overline{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{nm}{n} = m \quad \text{où } m = E(X)$$

$$V(\overline{X}) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \to 0 \quad \text{lorsque } n \to +\infty$$

Donc d'après Tchebychev $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i \xrightarrow{P} m = E(X)$ quand $n \to +\infty$

C'est la loi des grands nombres.

$\overline{X}$ est un exemple d'estimateur de la moyenne $m = E(X)$

Définition On appelle variance empirique, la statistique :

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Proposition : $S^2 = \frac{1}{n}\sum_{i=1}^{n}X_i^{\,2} - (\overline{X})^2$

Démo : $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 2X_i\overline{X} + \overline{X}^2)$

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 2X_i\overline{X} + \overline{X}^2) = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - 2\overline{X}\frac{1}{n}\sum_{i=1}^{n}X_i + \frac{n}{n}\overline{X}^2$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - 2\overline{X}^2 + \overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - (\overline{X})^2$$

Montrons que $S^2 \xrightarrow{P} \sigma^2$ lorsque $n \to +\infty$

D'après la loi des grands nombres, on a :

$\overline{X} = \frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{P} m = E(X)$ quand $n \to +\infty$

Et $\frac{1}{n}\sum_{i=1}^{n}X_i^{\,2} \xrightarrow{P} E(X^2)$ quand $n \to +\infty$

Donc $S^2 = \frac{1}{n}\sum_{i=1}^{n}X_i^{\,2} - (\overline{X})^2 \xrightarrow{P} E(X^2) - E^2(X) = \sigma^2 = V(X)$

S^2 est un estimateur de la variance .

Définition : On considère une population X, distribuée suivant une Loi de probabilité qui dépend d'un paramètre θ inconnu.

On prélève un échantillon $(X_1, X_2, \ldots\ldots, X_n)$ de X, on appelle estimateur de θ, toute variable aléatoire T_n fonction de l'échantillon :

$$T_n = f(X_1, X_2, \ldots., X_n)$$

On appelle biais de l'estimateur la quantité $b(T_n) = E(T_n) - \theta$

On dit que l'estimateur est sans biais si $b(T_n)=0 \Leftrightarrow E(T_n)=\theta$

Comme exemple $\overline{X}$ est un estimateur sans biais de $m=E(X)$

Puisque $E(\overline{X})=m$

<u>Définition :</u> On dit qu'une suite (T_n) d'estimateurs de θ est asymptotiquement sans biais si $\lim_{n\to+\infty} E(T_n)=\theta$

On appelle risque quadratique de T_n ou erreur quadratique :

$R(T_n)=E((T_n-\theta)^2)$

<u>Proposition :</u> le risque quadratique est $R(T_n)=V(T_n)+(E(T_n)-\theta)^2$

Démonstration :

$(T_n-\theta)^2=(T_n-E(T_n)+E(T_n)-\theta)^2$

$$E((T_n-\theta)^2)=E((T_n-E(T_n))^2)+2E((T_n-E(T_n))(E(T_n)-\theta))+E((E(T_n)-\theta)^2)$$

$$E((T_n-\theta)^2)=V(T_n)+2(E(T_n)-\theta)(E(T_n)-E(T_n))+(E(T_n)-\theta)^2$$

Donc $R(T_n)=V(T_n)+(E(T_n)-\theta)^2$

Remarque : Si l'estimateur est sans biais $b(T_n)=E(T_n)-\theta=0$

Alors $R(T_n)=V(T_n)$

Donc si on a deux estimateurs sans biais du paramètre θ

Le plus précis est celui de variance minimale.

<u>Définition :</u> On dit que l'estimateur T_n est convergent si cet Estimateur converge en probabilité vers le paramètre θ.

on écrira $T_n \xrightarrow{P} \theta$ lorsque $n\to+\infty$

<u>Définition :</u> On appelle vraisemblance de θ , la densité de L'échantillon $(X_1,X_2,.....,X_n)$:

$$L(x_1, x_2, ..., x_n, \theta) = \prod_{i=1}^{n} P(X_i = x_i) \text{ (dans le cas discret)}$$

$$L(x_1, x_2, ..., x_n, \theta) = \prod_{i=1}^{n} f(x_i) \text{ (dans le cas continu)}$$

<u>Exemple :</u> On considère un échantillon $(X_1, X_2, ..., X_n)$ d'une variable de Poisson de paramètre θ (inconnu)

la vraisemblance de cet échantillon est : $L(x_1, x_2, ..., x_n, \theta) = \prod_{i=1}^{n} P(X_i = x_i)$

$$L(x_1, x_2, ..., x_n, \theta) = \prod_{i=1}^{n} \exp(-\theta) \frac{\theta^{x_i}}{x_i!} = \frac{\exp(-n\theta)\theta^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

<u>Définition :</u> On appelle quantité d'information de Fisher $I_n(\theta)$ apportée par Un échantillon sur le paramètre θ la quantité positive : $I_n(\theta) = E((\frac{\partial \ln L}{\partial \theta})^2)$

<u>Proposition</u> :

$$I_n(\theta) = -E(\frac{\partial^2 \ln L}{\partial \theta^2})$$

Démo : L étant une densité : $\int_{IR^n} L(x, \theta) dx = 1$

En dérivant par rapport à θ : $\int_{IR^n} \frac{\partial L(x, \theta)}{\partial \theta} dx = 0$ (1)

en remarquant que $\frac{\partial \ln L(x, \theta)}{\partial \theta} = \frac{\frac{\partial L}{\partial \theta}(x, \theta)}{L(x, \theta)}$

(1) donne $\int_{IR^n} \frac{\partial \ln L(x, \theta)}{\partial \theta} L(x, \theta) dx = 0$

ce qui prouve que la variable aléatoire $\frac{\partial \ln L(x, \theta)}{\partial \theta}$ est centrée et que $I_n(\theta) = V(\frac{\partial \ln L}{\partial \theta})$

Dérivons une deuxième fois par rapport à θ :

$$\int_{IR^n} \frac{\partial^2 \ln L(x,\theta)}{\partial\theta^2} L(x,\theta)dx + \int_{IR^n} \frac{\partial \ln L(x,\theta)}{\partial\theta} \frac{\partial L(x,\theta)}{\partial\theta} dx = 0$$

$$\int_{IR^n} \frac{\partial^2 \ln L(x,\theta)}{\partial\theta^2} L(x,\theta)dx + \int_{IR^n} \left(\frac{\partial \ln L(x,\theta)}{\partial\theta}\right)^2 L(x,\theta)dx = 0$$

Donc

$$I_n(\theta) = E\left(\left(\frac{\partial \ln L}{\partial\theta}\right)^2\right) = \int_{IR^n} \left(\frac{\partial \ln L(x,\theta)}{\partial\theta}\right)^2 L(x,\theta)dx = -\int_{IR^n} \frac{\partial^2 \ln L(x,\theta)}{\partial\theta^2} L(x,\theta)dx = -E\left(\frac{\partial^2 \ln L(x,\theta)}{\partial\theta^2}\right)$$

Inégalité de FRECHET-DARMOIS-CRAMER-RAO(FDCR)

On a pour tout estimateur T sans biais de θ : $V(T) \geq \frac{1}{I_n(\theta)}$

L'estimateur T sera qualifié d'efficace si la borne inférieure est atteinte

C'est-à-dire $V(T) = \frac{1}{I_n(\theta)}$

Méthode du maximum de vraisemblance

Cette méthode consiste, étant donnée un échantillon de valeurs $x_1, x_2, ..., x_n$ à prendre comme estimation de θ la valeur de θ qui rend maximale la vraisemblance $L(x_1, x_2, ..., x_n, \theta)$.

En fait, on prend comme estimation de θ la solution de l'équation de la vraisemblance : $\frac{\partial \ln L}{\partial\theta} = 0$