

STABILISATION DE LA VARIANCE (C2, J4)

(07 / 10 / 2019, © Monfort, Dicostat2005, 2005-2019)

(i) Soit $(\mathcal{X}, \mathcal{B}, (P_\theta^X)_{\theta \in \Theta})$ un **modèle statistique** image dans lequel $\mathcal{X} = \mathbf{R}^N$ et $\Theta \subset \mathbf{R}^Q$. Etant donné une fonction $g : \Theta \mapsto \mathbf{R}$, on considère une **application mesurable** $t_N : \mathcal{X} \mapsto \mathbf{R}$ et un **estimateur** $T_N = t_N(X)$ du **paramètre** transformé $\tau = g(\theta)$. On suppose que :

$$(a) T_N \in \mathcal{L}^2(\Omega, \mathcal{F}, P_\theta), \forall \theta \in \Theta ;$$

(b) $E_\theta T_N = \tau$ (**estimateur sans biais**). On écrit alors $V_\theta T_N = v_N^2(\tau)$ pour exprimer que la **variance** de T_N dépend de τ .

(ii) En général, il n'existe pas de transformation $\varphi : \mathbf{R} \mapsto \mathbf{R}$ tq $V_\theta \varphi(T_N) = \gamma$ soit une **constante** indépendant de τ .

Par contre, il peut exister une application $\varphi : \mathbf{R} \mapsto \mathbf{R}$ tq :

$$(1) \quad V_\theta \varphi(T_N) = \gamma \cdot \{1 + O(N^{-1})\}$$

(où O désigne la notation « grand zéro »).

Cette application φ est appelée **transformation stabilisant la variance**, et sa mise en oeuvre s'appelle **stabilisation de la variance**.

On montre qu'elle est donnée par l'**intégrale indéfinie** :

$$(2) \quad \varphi(u) = \alpha \cdot \int \{v_N(u)\}^{-1} du, \text{ avec } \alpha \in \mathbf{R}.$$

(iii) Par ailleurs, on appelle **stabilisation de la variance** l'usage d'une transformation $\varphi : \mathbf{R} \mapsto \mathbf{R}$ tq :

$$(3) \quad D \varphi(\tau) = (I_\theta)^{1/2},$$

où $I_\theta = E_\theta \{D_2 \text{Log } f(X, \theta)\}^2$ est l'**information de FISHER**, avec $f(x, \theta) = dP_\theta^X / d\mu(x)$ (**vraisemblance** du modèle), où μ est une **mesure positive** dominante σ -finie et $\theta \in \mathbf{R}$.

Cette définition est utilisée en conjonction avec :

(a) la propriété d'**invariance du maximum de vraisemblance** : ie si $\tilde{\tau}$ est estimateur du mv de $\tau = g(\theta)$, alors $\varphi(\tilde{\tau})$ est un estimateur du mv de $\varphi(\tau)$;

(b) la **normalité asymptotique** de l'estimateur précédent, ie :

$$(4) \quad N^{1/2} \{\varphi(\tilde{\tau}) - \varphi(\tau)\} \xrightarrow{\mathcal{L}_{N \rightarrow +\infty}} \mathcal{N}(0, 1) \text{ (loi normale réduite),}$$

où l'**échantillon** $X = (X_1, \dots, X_N)$ est supposé indépendamment distribué (cf **échantillon indépendant**).

(iv) La stabilisation de la variance intervient souvent dans l'étude du **modèle de régression** ou du **modèle d'analyse de la variance** (notamment ceux associés à un **plan d'expérience**) : ainsi en est-il lorsque la **perturbation aléatoire** de ces modèles n'est pas homoscédastique (cf **homoscédasticité**).