2 Cas général : un n-échantillon d'une loi de moyenne $m = \mathbb{E}_\theta[X_1]$ et $\sigma^2 = V_\theta(X_1)$.

- Comme S_n^2 ne dépend pas de m on va supposer que $m = 0$ et alors $\sigma^2 = \mathbb{E}_\theta[X_1{}^2]$. Alors

$$\sqrt{n}[S_n^2 - \sigma^2] = \sqrt{n}\left[\frac{1}{n}\sum_i X_i{}^2 - (\overline{X}_n)^2 - \sigma^2\right] \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\left(0, V_\theta(X_1{}^2)\right)$$

– TCL : $\sqrt{n}[\overline{X^2}_n - \sigma^2] \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, V_\theta(X_1{}^2))$

– Slutsky : $\sqrt{n}(\overline{X}_n)^2 = \frac{1}{\sqrt{n}} \times n(\overline{X}_n)^2 \xrightarrow[n\to+\infty]{P} 0$ car $\sqrt{n}\,\overline{X}_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$ avec le TCL d'où $g(\overline{X}_n) = n(\overline{X}_n)^2 \xrightarrow[n\to+\infty]{\mathcal{L}} X^2$.

- Pour $m \neq 0$ alors on applique le résultat précédent aux variables d'espérance nulle : $X_i - m$

$$\sqrt{n}[S_n^2 - \sigma^2] \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\left(0, V_\theta((X_1 - m)^2)\right)$$

avec

$$\begin{aligned} V_\theta((X_1 - m)^2) &= \mathbb{E}_\theta[(X_1 - m)^4] - \mathbb{E}_\theta[(X_1 - m)^2]^2 \\ &= \mathbb{E}_\theta[(X_1 - m)^4] - V_\theta(X_1)^2 \end{aligned}$$

Le résultat suivant donne une évaluation **exacte** de la vitesse de convergence pour des n-échantillons d'une loi simple

Théorème 11 (Théorème de Berry-Esseen (voir [BJ06, Cor 11.64])

Pour un n-échantillon $(X_1, \dots, X_n)$ d'une loi commune F de moyenne m et de variance σ^2 et admettant un moment d'ordre 3, il existe une constante indépendante de F telle que,

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \sqrt{n}\left[\frac{\overline{X}_n - m}{\sigma}\right] \le t \right\} - \Phi(t) \right| \le \frac{C}{\sqrt{n}} \frac{\mathbb{E}|X_1 - m|^3}{\sigma^3} \tag{27}$$

Remarque 20

- Cette inégalité est valable pour toute taille n de l'échantillon et quelque soit la loi commune sous-jacente.
- La constante C optimale n'est pas connue. Cependant, elle peut être prise égale à 0.7975 mais n'est pas plus petite que 0.4097.
- L'élément le plus important dans (27) est le fait que la constante soit indépendante de la loi commune de l'échantillon.

La loi commune de l'échantillon dépend de la taille

On dispose d'un n-échantillon d'une loi de Poisson de paramètre $m = 1/n$, i.e. la loi commune dépend de la taille n de l'échantillon. Alors la version du TCL du Th 10 ne s'applique pas. Par ailleurs : $S_n \sim \text{Pois}(1)$ donc de moyenne 1, de variance 1 et la loi de la v.a.

$$\sqrt{n}\left[\frac{\overline{X}_n - m_n}{\sigma_n}\right] = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{V(S_n)}} = S_n - 1$$

ne dépend pas de n et ne converge pas vers $\mathcal{N}(0,1)$

Corollaire 4

Sous les hypothèses du Th 11, on a

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left\{ \sqrt{n}\left[\frac{\overline{X}_n - m}{\sigma}\right] \le t \right\} - \Phi(t) \right| \xrightarrow[n\to+\infty]{} 0$$

pour toute loi commune F_n du n-échantillon, de moyenne m_n et σ_n^2 vérifiant

$$\frac{\mathbb{E}_n|X_1 - m_n|^3}{\sigma_n^3} = o(\sqrt{n}).$$

Ici $\mathbb{E}_n$ est l'espérance sous F_n.

Loi binomiale avec p_n

- Si $S_n \sim \text{Bin}(n,p)$ alors

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

- On considère le cas où $p \equiv p_n$ dépend de n. Alors, comme S_n est la somme d'un n-échantillon d'une $\text{Ber}(p_n)$ alors le Cor 4 nous donne

$$\frac{S_n - np_n}{\sqrt{np_n(1-p_n)}} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

pourvu que

$$\gamma_n = \frac{p_n{}^2 + (1-p_n)^2}{\sqrt{p_n(1-p_n)}} = o(\sqrt{n})$$

  - cette condition est satisfaite si $p_n \xrightarrow[n\to+\infty]{} p$ avec $0 < p < 1$
  - si $p_n \xrightarrow[n\to+\infty]{} 0$ alors $\gamma_n \sim 1/\sqrt{p_n}$ et une condition suffisante pour la convergence vers la loi normale est

$$\frac{1}{\sqrt{p_n}} = o(\sqrt{n}) \quad \text{ou} \quad \frac{1}{n} = o(p_n)$$

Notons que si $p_n = 1/n$ alors on se retrouve dans le cadre de la convergence en loi vers une loi de Poisson

- La borne de l'erreur d'approximation (27) est valide pour toute loi commune de l'échantillon (et indépendante de la taille). On peut espérer une meilleure borne pour une loi spécifique. Le résultat suivant donne la réponse

Théorème 12 (Développement d'Edgeworth)

Pour un n-échantillon d'une loi F qui n'est pas arithmétique et admet un moment d'ordre 3 alors

$$(28)\quad \mathbb{P}\left\{\sqrt{n}\left[\frac{\overline{X}_n - m}{\sigma}\right] \le t\right\} = \Phi(t) + \frac{\mu_3}{6\sigma^3\sqrt{n}}(1-t^2)\exp(-t^2/2) + o\left(\frac{1}{\sqrt{n}}\right)$$

où μ_3 est le moment centré d'ordre 3 de la loi F. Le second terme dans le membre de droite est la correction d'Edgeworth d'ordre 1 de l'approximation gaussienne.

Remarque 21

- En fait, il peut être montré que le terme d'erreur $o(1/\sqrt{n})$ est uniforme en t.
- La correction est nulle dès que $\mu_3 = 0$, par exemple pour toute loi symétrique autour de m, et la convergence est alors plus rapide que $1/\sqrt{n}$. Par contre si $\mu_3 \neq 0$, alors la convergence est en $1/\sqrt{n}$.

Correction de continuité à la Yates

- Dans le théorème précédent, la loi commune de l'échantillon est supposée non-arithmétique. En effet, si cela n'est pas le cas alors le développement (28) n'est pas correct car un autre terme d'ordre $1/\sqrt{n}$ doit être pris en compte (voir ci-dessous), appelé le **terme de correction de la continuité**.
- En effet, donnons un support intuitif à l'adjonction de terme supplémentaire : supposons que la différence ΔX_1 entre deux valeurs successives de X_1 soit h alors $\Delta Z_n = \frac{h}{\sigma\sqrt{n}}$. « L'idée » est d'utiliser l'approximation suivant de $\mathbb{P}\{Z_n = z\}$ en termes d'aire sous la courbe de la densité de la loi normale

$$\mathbb{P}\{Z_n = z\} \approx \Phi\left(z + \frac{h}{2\sigma\sqrt{n}}\right) - \Phi\left(z - \frac{h}{2\sigma\sqrt{n}}\right)$$

$$\text{ou } \mathbb{P}\{Z_n \le z\} \approx \Phi\left(z + \frac{h}{2\sigma\sqrt{n}}\right)$$

- Ensuite, un développement de Taylor nous donne que

$$\Phi\left(z + \frac{h}{2\sigma\sqrt{n}}\right) = \Phi(z) + \frac{h}{2\sigma\sqrt{n}}\exp(-z^2/2) + o\left(\frac{1}{\sqrt{n}}\right)$$

« d'où » une version du développement (28) dans le cas arithmétique :

$$\mathbb{P}\left\{\sqrt{n}\left[\frac{\overline{X}_n - m}{\sigma}\right] \le t\right\} = \Phi(t) + \frac{\exp(-t^2/2)}{\sqrt{n}}\left[\frac{\mu_3}{6\sigma^3}(1-t^2) + \frac{h}{2\sigma}\right] + o\left(\frac{1}{\sqrt{n}}\right)$$

Normalité asymptotique d'un estimateur

- Comme pour les diverses formes de consistance, un estimateur $\widehat{\theta}_n$ sera dit convergent en loi vers une v.a. Z si cette convergence a lieu pour **toute valeur du paramètre** $\theta \in \Theta$.
  - Autrement dit la suite $\{\widehat{\theta}_n\}_{n\ge 1}$ doit satisfaire l'une des trois conditions équivalentes de la définition de convergence en loi pour tout $\theta \in \Theta$. Par exemple : pour toute fonction continue bornée f à valeurs réelles,

$$\forall \theta \in \Theta, \quad \mathbb{E}_\theta[f(\widehat{\theta}_n)] \xrightarrow[n\to+\infty]{} \mathbb{E}_\theta[f(Z)]$$

  Ceci étant, cela ne signifie pas une convergence uniforme en le paramètre.

- Nous allons introduire la **propriété de normalité asymptotique d'un estimateur** consistant d'un paramètre. Elle sera au cœur des constructions d'intervalles de confiance et des tests lorsque les échantillons ne seront pas gaussiens.

Définition 23 (Estimateur asymptotiquement normal ou gaussien)

Un estimateur $\widehat{\theta}_n$ d'un paramètre $\theta \in \Theta$, défini sur un certain modèle statistique d'échantillonnage, est dit asymptotiquement normal ou possédé la propriété de normalité asymptotique, si

$$(29)\qquad \sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, \Sigma(\theta))$$

$\Sigma(\theta)$ est appelée la variance asymptotique de l'estimateur.

- Notons que tout estimateur de θ asymptotiquement normal est nécessairement consistant, et que cette dernière propriété est équivalente à la convergence de $\{\widehat{\theta}_n\}_{n\ge 1}$ en loi vers θ.

$$\widehat{\theta}_n \xrightarrow[n\to+\infty]{P} \theta \iff \widehat{\theta}_n \xrightarrow[n\to+\infty]{\mathcal{L}} \theta.$$

- Notons que la définition s'applique à tout estimateur de $g(\theta)$.

Définition 24 (Estimateur asymptotiquement efficace)

Plaçons nous dans le cadre d'échantillonnage d'un modèle statistique pour une observation homogène et dans lequel l'information de Fisher $I(\theta)$ est inversible. Un estimateur asymptotiquement normal $\widehat{\theta}_n$ du paramètre $\theta \in \Theta$, est dit asymptotiquement efficace si la variance asymptotique $\Sigma(\theta)$ est égale à l'inverse de l'information de Fisher

(30) $$\Sigma(\theta) = I(\theta)^{-1}.$$

$I(\theta)^{-1}$ est souvent appelée la borne de FDRC asymptotique.

- Dans le cas d'un estimateur $\widehat{\theta}_n$ de $g(\theta)$ avec g dérivable sur un ouvert Θ, alors (30) devient
$$\Sigma(g(\theta)) = g'(\theta)^2 \, I(\theta)^{-1}$$
- Sous certaines conditions, on peut montrer que la variance asymptotique vérifie $\Sigma(\theta) \geq I(\theta)^{-1}$. Un estimateur asymptotiquement efficace est donc un estimateur dont la variance asymptotique est égale à la borne de FDRC asymptotique.

« Formule de Taylor » et la delta-méthode

- On souhaite déduire la normalité asymptotique pour un estimateur $g(\widehat{\theta}_n)$ fonction d'un estimateur $\widehat{\theta}_n$ de θ asymptotiquement normal.
- Pour cela, si g est une fonction suffisamment régulière, une idée naturelle est d'exploiter un développement de Taylor pour ramener l'étude du comportement de $g(X_n)$ à celle de X_n.

Théorème 13

Si

(31) $$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, \Sigma)$$

alors

(32) $$\sqrt{n}(g(\widehat{\theta}_n) - g(\theta)) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, \Sigma \times g'(\theta)^2)$$

pourvu que la dérivée $g'(\theta)$ existe et soit non nulle.

Modèle de Poisson $\mathrm{Pois}(\theta)$

- $\overline{X}_n$ est un ESBVM de θ et le TCL donne
$$\sqrt{n}\left[\frac{\widehat{\theta}_n - \theta}{\sqrt{\theta}}\right] \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$$
et donc $\overline{X}_n$ est asymptotiquement normal
$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,\theta)$$
- $g(\theta) = \exp(-\theta)$ avec $g'(\theta) = -\exp(-\theta)$: alors l'EMV $g(\overline{X}_n)$ est biaisé et asymptotiquement normal
$$\sqrt{n}[g(\widehat{\theta}_n) - g(\theta)] \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, (-\exp(-\theta))^2 \times \theta)$$

Technique de stabilisation de variance

Dans la propriété de normalité asymptotique,

(33) $$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, \Sigma(\theta))$$

la variance asymptotique $\Sigma(\theta)$ dépend généralement de θ. Il peut être utile pour la construction d'intervalle de confiance par exemple, de déterminer un autre estimateur asymptotiquement normal de variance asymptotique **indépendante** de θ.

- Supposons que (33) est valide pour l'estimateur $\widehat{\theta}_n$. Alors la méthode delta permet d'écrire que

(34) $$\sqrt{n}(g(\widehat{\theta}_n) - g(\theta)) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, \Sigma(\theta) \times g'(\theta)^2)$$

pourvu que la dérivée $g'(\theta) \neq 0$ existe. La loi limite sera de variance V^2 indépendante de θ si

(35) $$g'(\theta) = \frac{V}{\sqrt{\Sigma(\theta)}}$$

Un exemple

On vient de voir pour le modèle de Poisson que $\widehat{\theta}_n = \overline{X}_n$ est asymptotiquement normal avec $\Sigma(\theta) = \theta$. Alors l'équation (35) s'écrit

$$g'(\theta) = \frac{V}{\sqrt{\theta}}$$

dont une solution est $g(\theta) = 2V\sqrt{\theta}$. En prenant $V = 1$ alors $g(\theta) = 2\sqrt{\theta}$ et on obtient

$$2\sqrt{n}(\sqrt{\widehat{\theta}_n} - \sqrt{\theta}) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

Extension de la méthode delta dans le cas où $g'(\theta) = 0$

Utiliser un développement de Taylor à un ordre supérieur si la régularité de la fonction g est suffisante. Supposons que $g''(\theta) \neq 0$ alors un dév. à l'ordre 2 donne

$$g(\widehat{\theta}_n) = g(\theta) + \underbrace{(\widehat{\theta}_n - \theta)g'(\theta)}_{=0} + \frac{1}{2}(\widehat{\theta}_n - \theta)^2 g''(\theta) + (\widehat{\theta}_n - \theta)^2 o(1)$$

Alors

$$k_n\left[g(\widehat{\theta}_n) - g(\theta)\right] = \frac{k_n}{2}(\widehat{\theta}_n - \theta)^2 g''(\theta) + \frac{k_n}{2}(\widehat{\theta}_n - \theta)^2 o(1)$$

Maintenant si $\widehat{\theta}_n$ est asymptotiquement normal de matrice $\Sigma(\theta)$ alors la Prop 15-2. nous donne

$$\left(\sqrt{n}\frac{\widehat{\theta}_n - \theta}{\sqrt{\Sigma(\theta)}}\right)^2 \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)^2 = \chi^2_1$$

Autrement dit, la même preuve que celle donnée du Th 13 donnera

$$n[g(\widehat{\theta}_n) - g(\theta)] \xrightarrow[n\to+\infty]{\mathcal{L}} \frac{\Sigma(\theta)g''(\theta)}{2}\chi^2_1$$

Modèle d'échantillon d'une $\mathrm{Ber}(\theta)$

Le TCL nous donne la normalité asymptotique de l'estimateur $\widehat{\theta}_n = \overline{X}_n$ de θ :

$$\sqrt{n}(\overline{X}_n - \theta) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, \Sigma(\theta) = \theta(1-\theta))$$

L'EMV de $g(\theta) = \theta(1-\theta)$ est $\widehat{\theta}_n = \overline{X}_n(1 - \overline{X}_n)$ alors la méthode delta avec $g'(\theta) = 1 - 2\theta \neq 0$, i.e. $\theta \neq 1/2$:

$$\sqrt{n}[g(\widehat{\theta}_n) - g(\theta)] \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\left(0, (1-2\theta)^2 \times \theta(1-\theta)\right)$$

Maintenant si $\theta = 1/2$ alors $g'(\theta) = 0$ et $g''(\theta) = -2$ alors

$$n[g(\widehat{\theta}_n) - 1/4] \xrightarrow[n\to+\infty]{\mathcal{L}} -\frac{1}{4}\chi^2_1$$

Convergence en loi : cas des vecteurs aléatoires

Théorème et définition (cf cours de proba.)

Une suite $\{\underline{X}_n = (X_n(1), \ldots, X_n(p))\}_{n\geq 0}$ à valeurs dans $\mathbb{R}^p$ définis sur un modèle statistique est dite convergente en loi vers un vecteur aléatoire $\underline{X}$ ssi l'une des 2 conditions suivantes est réalisée

1. pour toute fonction continue bornée f sur $\mathbb{R}^p$,

$$\mathbb{E}[f(\underline{X}_n)] \xrightarrow[n\to+\infty]{} \mathbb{E}[f(\underline{X})]$$

2. la suite des fonctions caractéristiques $\{\varphi_{\underline{X}_n}\}_{n\geq 0}$ converge en tout point de $\mathbb{R}$ vers φ_Z :

$$\forall t \in \mathbb{R}^p, \quad \varphi_{\underline{X}_n}(t) := \mathbb{E}[\exp(i\langle t, \underline{X}_n\rangle)] \xrightarrow[n\to+\infty]{} \varphi_{\underline{X}}(t)$$

Remarque 22

- Nous avons repris les deux critères de convergence les plus utilisés. Cependant, le critère sur les fonctions de répartition est également valable.
- Il est clair que la convergence d'une suite de vecteurs aléatoires entraîne la convergence composante par composante.

- À partir du critère de convergence par les fonctions caractéristiques, il est facile d'en déduire un critère ramenant la convergence de vecteurs aléatoires à celle de v.a. Cette « réduction » a été proposée par Cramér et Wold

Théorème 14 (Réduction de Cramér et Wold)

$\underline{X}_n \xrightarrow[n\to+\infty]{\mathcal{L}} \underline{X}$ ssi pour toute famille de constantes $(c(1),\dots,c(p)) \in \mathbb{R}^p$

$$\sum_{i=1}^{p} c(i)X_n(i) \xrightarrow[n\to+\infty]{\mathcal{L}} \sum_{i=1}^{p} c(i)X(i)$$

Une conséquence directe de la réduction de Cramér et Wold est la version multidimensionnelle du TCL.

Théorème 15 (TCL multidimensionnel)

Considérerons une suite i.i.d. de vecteurs aléatoires de $\mathbb{R}^p$ de moyenne $\underline{m}$ et de matrice de variance-covariance Σ. Alors

$$\sqrt{n}\Big[\frac{1}{n}\sum_{k=1}^{n}\underline{X}_k - \underline{m}\Big] \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,\Sigma)$$

où $\mathcal{N}(0,\Sigma)$ représente un vecteur gaussien centré et de matrice de variance $\Sigma = \mathrm{Cov}(\underline{X}_1)$. Cette dernière matrice est appelée la matrice de variance asymptotique.

Remarque 23

Si on souhaite rester dans un cadre dominé, alors on rappelle qu'un vecteur gaussien (limite) admet une densité ssi la matrice $\Sigma \in \mathcal{L}^{++}$.

- La définition 23 d'estimateur asymptotiquement normal se transpose directement au cas des vecteurs aléatoires. De même pour la définition 24 d'efficacité asymptotique.

Normalité asymptotique des moments empiriques non centrés

- L'estimateur $\widehat{\theta}_n = (\overline{X}_n, \overline{X^2}_n)$ est un estimateur consistant de m_1, m_2.
- Prouvons sa normalité asymptotique : notons que

$$(\overline{X}_n, \overline{X^2}_n) = \frac{1}{n}\sum_{i=1}^{n}\underline{X}_i \tag{36}$$

où $\underline{X}_i = (X_i, X_i{}^2)$ est la moyenne empirique d'une suite de vecteurs i.i.d. Sous des conditions d'existence de moment d'ordre 4, le TCL nous donne

$$\sqrt{n}\begin{pmatrix}\overline{X}_n - m_1\\ \overline{X^2}_n - m_2\end{pmatrix} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\left(0, \begin{pmatrix} m_2-(m_1)^2 & m_3-m_1m_2\\ m_3-m_1m_2 & m_4-(m_2)^2\end{pmatrix}\right)$$

- Le résultat se généralise pour l'estimateur des k premiers moments empiriques

- **Le résultat du second point de la Prop 20 sur la transformation d'une suite convergente en loi par une fonction continue sur un ensemble de points de probabilité 1 relativement à la loi limite, est également valable pour des fonctions (multidimensionnelles) de vecteurs aléatoires** d'après le même résultat pour la convergence ps (cf la remarque qui suit la Prop 15).

Corollaire 5

- *Si $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$ avec X_n vecteur aléatoire de dimension k alors pour toutes matrices $A \in M_{p,k}(\mathbb{R})$ et $B \in M_{k,k}(\mathbb{R})$,*

$$AX_n \xrightarrow[n\to+\infty]{\mathcal{L}} AX \quad \text{et} \quad X_n{}^\top BX_n \xrightarrow[n\to+\infty]{\mathcal{L}} X^\top BX$$

- *Si $\{X_n\}_{n\geq 0}$ est une suite de matrices aléatoires inversibles qui convergent en loi (consistante) vers une matrice inversible X alors $\{(X_n)^{-1}\}_{n\geq 0}$ converge vers X^{-1}.*

Différence de deux moyennes

On dispose de deux échantillons l'un, $(X_1,\dots,X_n)$ de la loi F de moyenne m_F et variance σ_F^2, l'autre $(Y_1,\dots,Y_m)$ de la loi G de moyenne m_G et variance σ_G^2. On suppose que les deux échantillons sont également indépendants. Alors le TCL

$$\sqrt{n}[\overline{X}_n - m_F] \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,\sigma_F^2) \quad \sqrt{m}[\overline{Y}_m - m_G] \xrightarrow[m\to+\infty]{\mathcal{L}} \mathcal{N}(0,\sigma_G^2)$$

On pose $n=(1-p)N$ et $m=pN$ où $0<p<1$ alors $N=n+m$ et par Slutsky

$$\frac{\sqrt{N}}{\sqrt{n}}\sqrt{n}[\overline{X}_n - m_F] \xrightarrow[N\to+\infty]{\mathcal{L}} \mathcal{N}(0,\frac{\sigma_F^2}{1-p}) \quad \sqrt{N}[\overline{Y}_m - m_G] \xrightarrow[N\to+\infty]{\mathcal{L}} \mathcal{N}(0,\frac{\sigma_G^2}{p})$$

Par l'indépendance des deux échantillons, on a

$$\left(\sqrt{N}[\overline{X}_n - m_F], \sqrt{N}[\overline{Y}_m - m_G]\right) \xrightarrow[N\to+\infty]{\mathcal{L}} (X,Y)$$

où (X,Y) sont deux v.a. indépendantes de loi $\mathcal{N}(0,\frac{\sigma_F^2}{1-p})$ et $\mathcal{N}(0,\frac{\sigma_G^2}{p})$ respectivement. Maintenant $X-Y=g(X,Y)$ est continue et on obtient

$$\sqrt{N}\left((\overline{X}_n - \overline{Y}_m) - (m_F - m_G)\right) \xrightarrow[N\to+\infty]{\mathcal{L}} \mathcal{N}(0,\frac{\sigma_F^2}{1-p}+\frac{\sigma_G^2}{p}) \tag{37}$$

ou encore

$$\frac{(\overline{X}_n - \overline{Y}_m) - (m_F - m_G)}{\sqrt{\frac{\sigma_F^2}{n}+\frac{\sigma_G^2}{m}}} \xrightarrow[N\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

- Le premier point de la Prop 20 est également valide modulo la reformulation dans le cadre multidimensionnel des expressions spécifiques au cas unidimensionnel.

Corollaire 6

Supposons que $\widehat{\theta}_n$ est asymptotiquement normal de $\theta\in\mathbb{R}^p$ de matrice de covariance asymptotique $\Sigma(\theta)\in\mathcal{L}^{++}$. Dans ce cas, on peut définir la matrice « racine carrée » $\Sigma(\theta)^{1/2}$ de $\Sigma(\theta)$, i.e. l'unique élément de $\mathcal{L}^{++}$ tel que $\Sigma(\theta)^{1/2}\Sigma(\theta)^{1/2}=\Sigma(\theta)$.
Si $\widehat{\Sigma}_n\in\mathcal{L}^{++}$ est un estimateur consistant de cette matrice de covariance asymptotique, c'est à dire

$$\widehat{\Sigma}_n \xrightarrow[n\to+\infty]{P} \Sigma(\theta)$$

alors

- $\sqrt{n}\,\widehat{\Sigma}_n^{-1/2}(\widehat{\theta}_n-\theta) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,I_p)$
- $n(\widehat{\theta}_n-\theta)^\top\widehat{\Sigma}_n^{-1}(\widehat{\theta}_n-\theta) \xrightarrow[n\to+\infty]{\mathcal{L}} \chi_p^2$

- **Méthode Delta**

Théorème 16

1. *Considérons une fonction $g:\theta\in\mathbb{R}^p\to g(\theta)\in\mathbb{R}^q$ différentiable telle que la matrice $q\times p$ $J(g)=\left(\frac{\partial g_i}{\partial\theta_j}(\theta)\right)_{i\in\mathbb{R}^q;j\in\mathbb{R}^p}$ soit inversible en θ*
2. *$\widehat{\theta}_n$ est un estimateur asymptotiquement normal de θ, de matrice de covariance asymptotique $\Sigma(\theta)$.*

Alors, $g(\widehat{\theta}_n)$ est un estimateur asymptotiquement normal de $g(\theta)$ de matrice de covariance asymptotique

$$J(g)\Sigma(\theta)J(g)^\top \tag{38}$$

Corollaire 7 (Convergence pour g à valeurs réelles)

Sous les conditions du précédent théorème, $g(\widehat{\theta}_n)$ est un estimateur asymptotiquement normal de $g(\theta)$ de matrice de covariance asymptotique $\Sigma(g(\theta))$

$$\Sigma(g(\theta))=\sum_{i=1}^p\sum_{j=1}^p\frac{\partial g}{\partial\theta_i}(\theta)\Sigma(\theta)_{i,j}\frac{\partial g}{\partial\theta_j}(\theta)$$

Normalité asymptotique de la variance empirique revisitée

Cas d'un n-échantillon d'une loi de moyenne $\mathbb{E}_\theta[X_1]=m$ et de variance $V_\theta(X_1)=\sigma^2$.
$\theta=(m,\sigma^2)$ et on s'intéresse à l'estimateur S_n^2 de la variance σ^2

- On peut supposer que $m=0$ sans perte de généralité et $\theta=(0,m_2)$ avec l'estimateur $\widehat{\theta}_n=(\overline{X}_n,\overline{X^2}_n)$ et le TCL nous avait donné (39)
$$\sqrt{n}\left[\begin{pmatrix}\overline{X}_n\\ \overline{X^2}_n\end{pmatrix}-\begin{pmatrix}0\\ m_2\end{pmatrix}\xrightarrow[n\to+\infty]{\mathcal{L}}\mathcal{N}\left(0,\begin{pmatrix}m_2-(m_1)^2 & m_3-m_1m_2\\ m_3-m_1m_2 & m_4-(m_2)^2\end{pmatrix}\right)\right]$$
- $S_n^2=g(\overline{X}_n,\overline{X^2}_n)$ avec $g(\theta_1,\theta_2)=\theta_2-(\theta_1)^2$ avec
$$J(g)=\left(\frac{\partial g}{\partial\theta_j}(0,m_2)\right)=\begin{pmatrix}0 & 1\end{pmatrix}$$
- Le Cor 7 nous donne
$$\sqrt{n}[S_n^2-m_2]\xrightarrow[n\to+\infty]{\mathcal{L}}\mathcal{N}\left(0,J(g)\Sigma(\theta)J(g)^\top=\Sigma(\theta)_{2,2}=m_4-(m_2)^2\right)$$