où σ_n^2 est la variance empirique

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^2.$$

Il est alors clair que la fonction de vraisemblance atteint son maximum pour $\sigma^2 = \sigma_n^2$. σ_n^2 est donc ici l'estimateur du maximum de vraisemblance. C'est un estimateur consistant mais qui, comme on l'a déjà vu, est biaisé.

4. Dans le cas d'une loi gaussienne de paramètres (μ, σ^2), où σ^2 est inconnue et μ est inconnue, on refait comme précédemment. On a vu que

$$L(X_1, \dots, X_n, \mu, \sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2\right\}.$$

Le calcul fait au 3. montre qu'on a toujours

$$L(X_1, \dots, X_n, \mu, \sigma^2) \leq \frac{1}{(\sqrt{2\pi}\sigma_n)^n} \exp\left\{-\frac{1}{2\sigma_n^2} \sum_{i=1}^{n} (X_i - \mu)^2\right\},$$

avec égalité pour $\sigma = \sigma_n$. D'après le calcul fait en 2., ce majorant est maximal pour $\mu = \bar{X}_n$.

Finalement, le couple $\theta = (\mu, \sigma^2)$ réalisant le maximum de vraisemblance est formé par la moyenne empirique $\bar{X}_n$ et la variance empirique σ_n^2. Comme on l'a déjà signalé, $\bar{X}_n$ est consistant sans biais, tandis que σ_n^2 est consistant, mais biaisé. Pour avoir un estimateur sans biais de σ^2, il faut considérer l'estimateur $s_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$.

13.3.2 Méthode des moments

Soit un modèle statistique indexé par un ensemble Θ. Supposons donné un $n-$échantillon $X_1, \dots, X_n$ d'une loi intégrable sous $\mathbb{P}_\theta$ pour chaque $\theta \in \Theta$. On cherche un estimateur $\hat{\theta}$ de θ.

Supposons que Θ est une partie de $\mathbb{R}$. On suppose également qu'on a un intervalle I de $\mathbb{R}$ tel que $\mathbb{P}_\theta(X_1 \in I) = 1$ pour tout $\theta \in \Theta$. L'idée de la méthode des moments est de prendre comme estimateur $\hat{\theta}$ une valeur de θ telle que $\mathbb{E}_\theta X = m(\theta)$ coïncide avec la moyenne empirique observée.

Définition. Un estimateur par la méthode des moments est une solution de l'équation

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i = m(\hat{\theta}).$$

Si m réalise une bijection de Θ dans I , alors $\hat{\theta} = m^{-1}(\bar{X}_n)$.

Si Θ est une partie de $\mathbb{R}^d$, on calcule $\mathbb{E}_\theta X = m_1(\theta), \ldots, \mathbb{E}_\theta X^d = m_d(\theta)$. Un estimateur par la méthode des moments est une solution du système d'équations

$$\frac{1}{n}\sum_{i=1}^n X_i = m_1(\hat{\theta}), \quad \ldots\ , \quad \frac{1}{n}\sum_{i=1}^n X_i^d = m_d(\hat{\theta}).$$

Si m est inversible, alors $(\hat{\theta}_1, \ldots, \hat{\theta}_d)' = m^{-1}\left(\frac{1}{n}\sum\limits_{i=1}^n X_i, \ldots, \frac{1}{n}\sum\limits_{i=1}^n X_i^d\right)'$.

Exemple. On veut estimer la moyenne et la variance par la méthode des moments. Posons $\theta = (m, \sigma^2) \in \mathbb{R}\times]0, +\infty[$. On a $m = \mathbb{E}_\theta X$ et $\sigma^2 = \mathrm{Var}_\theta(X)$. On sait de plus que $\mathbb{E}_\theta(X^2) = \sigma^2 + m^2$, ce qui donne alors

$$\frac{1}{n}\sum_{i=1}^n X_i = \hat{m} \quad \text{et} \quad \frac{1}{n}\sum_{i=1}^n X_i^2 = \hat{\sigma}^2 + \hat{m}^2.$$

On en déduit donc que $\hat{m}$ est la moyenne empirique et

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - \left(\frac{1}{n}\sum_{i=1}^n X_i\right)^2.$$

Remarque.
— Dans certains cas, l'estimation par la méthode des moments est moins bonne que l'estimation par maximum de vraisemblance. Néanmoins, dans le cas de la loi Gamma par exemple, le calcul de la fonction de vraisemblance peut poser des problèmes (l'utilisation de l'ordinateur et d'algorithmes numériques est indispensable) tandis que l'estimation des moments est très facilement accessible.
— La méthode des moments peut s'utiliser comme point de départ pour maximiser la (log-)vraisemblance. En effet, on doit alors utiliser des algorithmes numériques, comme la Méthode de Newton, qui nécessitent des points de départ.
— Lorsque la taille de l'échantillon n'est pas suffisamment grande, la loi des grands nombres ne s'applique pas et par conséquent, les moments empiriques n'approchent pas suffisamment les moments théoriques. Ainsi, la méthode des moments n'est pas une bonne méthode dans ce cas, car les estimateurs obtenus peuvent sortir du support des paramètres. Par exemple, pour la loi $\Gamma(a)$, un petit échantillon peut conduire à $a < 0$.
— Enfin, nous avons vu que la méthode des moments consiste à résoudre une équation du type $\bar{X}_n = m(\hat{\theta})$, ce qui n'est pas toujours possible.