

Rappels : inférence statistique

Cem Ertur*

8 septembre 2014

— Version 2.0 —

1 Définitions

La statistique mathématique repose sur une idée simple : il existe un ensemble d'individus P appelé population, dont les caractéristiques ne sont pas connues. A partir de l'observation d'un sous-ensemble d'individus de cette population, appelé échantillon, on va chercher à déterminer, à **induire**, les principales caractéristiques de la population.

La statistique mathématique ou inférence statistique élabore des méthodes qui permettent de porter un jugement ou de décider au vu des résultats d'un échantillon.

Cette démarche inductive, qui va du particulier, l'échantillon, au général, la population, s'oppose au raisonnement déductif dans lequel on déduit à partir de quelques axiomes ou postulats, des théorèmes et des résultats en mathématiques, en microéconomie etc. En statistique le raisonnement procède par inférence ou induction : "opération mentale qui consiste à remonter des faits à la loi".

Le problème général de l'estimation se présente alors de la façon suivante : soit une population caractérisée par un paramètre inconnu, on considère un échantillon tiré de cette population. Connaissant la valeur du paramètre dans l'échantillon, on cherche à induire des informations sur la valeurs que peut prendre ce paramètre dans la population. On veut donc estimer les propriétés inconnues de la distribution de la population en se basant sur les propriétés connues de la distribution d'un échantillon tiré de cette population.

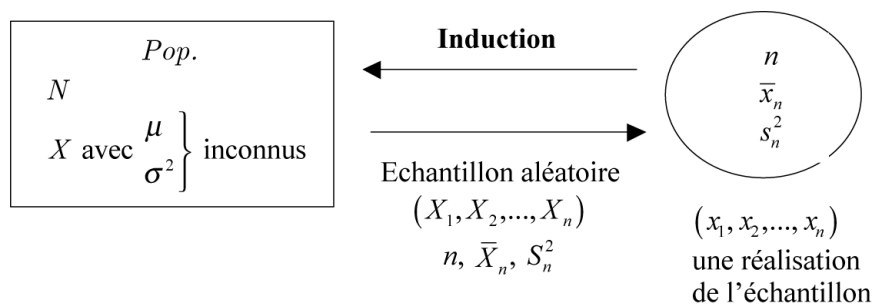


FIGURE 1 – population, échantillon aléatoire, induction

*Laboratoire d'Economie d'Orléans (UMR 6221 CNRS), UFR Droit Economie Gestion, Université d'Orléans. E-mail : cem.ertur@univ-orleans.fr, site internet : <http://certur.free.fr/>.

1.1 Echantillon aléatoire

Soit une population caractérisée par une variable aléatoire X .

Un échantillon aléatoire de taille n de X est une suite de n variables aléatoires (X_1, X_2, \dots, X_n) indépendantes et suivant toutes la même loi de probabilité que X , notée $f(X, \theta)$ où θ est un paramètre ou un vecteur de paramètres inconnus.

C'est une suite de variables aléatoires identiquement et indépendamment distribuées : *i.i.d.* de même distribution que la variable aléatoire X .

Les n valeurs (x_1, x_2, \dots, x_n) sont les valeurs observées ou la réalisation de l'échantillon aléatoire, des n variables aléatoires (X_1, X_2, \dots, X_n) .

1.2 Statistique

Etant donné un échantillon aléatoire, c'est-à-dire une suite de n variables aléatoires (X_1, X_2, \dots, X_n) identiquement et indépendamment distribuées, une statistique S_n est une fonction h des n variables aléatoires ; il s'agit donc aussi d'une variable aléatoire :

$$S_n = h(X_1, X_2, \dots, X_n) \quad (1)$$

La réalisation ou la valeur de la statistique correspond à une réalisation de l'échantillon aléatoire : c'est la valeur calculée de la statistique.

$$s_n = h(x_1, x_2, \dots, x_n)$$

Si un autre échantillon aléatoire était tiré sous des conditions identiques, on obtiendrait différentes valeurs observées, une réalisation différente de l'échantillon aléatoire et le valeur calculée de la statistique serait également différente présentant des fluctuations d'échantillonnage.

Exemple 1 Soit une population caractérisée par une variable aléatoire X . Supposons que dans la population $E(X) = \mu$ et $V(X) = \sigma^2$. Soit un échantillon aléatoire de taille n .

Soit \bar{X}_n la moyenne de l'échantillon : $\bar{X}_n = \sum_{i=1}^n X_i$

Soit S^{*2} la variance de l'échantillon : $S^{*2} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$

Calculons l'espérance mathématique et la variance de la moyenne de l'échantillon :

$$E(\bar{X}_n) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n \mu = \mu \quad (2)$$

$$V(\bar{X}_n) = V\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \underset{\text{par l'ind.}}{=} \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{1}{n^2} n \sigma^2 = \frac{\sigma^2}{n} \quad (3)$$

On constate que la variance de la moyenne de l'échantillon est toujours plus petite que la variance de la population : elle est inversement proportionnelle à la taille de l'échantillon, la dispersion de \bar{X}_n autour de μ est d'autant plus faible que la taille de l'échantillon est grande.

Si de plus, nous supposons que X est distribué suivant la loi Normale : $X \sim N(\mu, \sigma^2)$, les $X_i \sim i.i.d.N(\mu, \sigma^2)$ et on montre que $\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$ comme combinaison linéaire de variables aléatoires normales.

Calculons les caractéristiques de la variance de l'échantillon :

$$S_n^{*2} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{1}{n} \sum_{i=1}^n [(X_i - \mu) - (\bar{X}_n - \mu)]^2 \quad (4)$$

$$= \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 - \frac{2}{n} (\bar{X}_n - \mu) \sum_{i=1}^n (X_i - \mu) + (\bar{X}_n - \mu)^2 \quad (5)$$

or $\sum_{i=1}^n (X_i - \mu) = n(\bar{X}_n - \mu)$ donc on peut écrire :

$$S_n^{*2} = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 - (\bar{X}_n - \mu)^2$$

on peut maintenant calculer l'espérance mathématique de S^{*2} :

$$E(S_n^{*2}) = \frac{1}{n} \sum_{i=1}^n \underbrace{E[(X_i - \mu)^2]}_{=\sigma^2} - \underbrace{E[(\bar{X}_n - \mu)^2]}_{=V(\bar{X}_n)=\frac{\sigma^2}{n}} \quad (6)$$

$$E(S_n^{*2}) = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2 \quad (7)$$

Si nous avons posé $S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$, nous aurions obtenu le résultat suivant :
 $E(S^2) = \sigma^2$

En effet :

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{1}{n-1} \sum_{i=1}^n [(X_i - \mu) - (\bar{X}_n - \mu)]^2 \quad (8)$$

$$= \frac{1}{n-1} \sum_{i=1}^n (X_i - \mu)^2 - \frac{2}{n-1} (\bar{X}_n - \mu) \sum_{i=1}^n (X_i - \mu) + \frac{n}{n-1} (\bar{X}_n - \mu)^2 \quad (9)$$

$$= \frac{1}{n-1} \sum_{i=1}^n (X_i - \mu)^2 - \frac{2n}{n-1} (\bar{X}_n - \mu)^2 + \frac{n}{n-1} (\bar{X}_n - \mu)^2 \quad (10)$$

$$= \frac{1}{n-1} \sum_{i=1}^n (X_i - \mu)^2 - \frac{n}{n-1} (\bar{X}_n - \mu)^2 \quad (11)$$

l'espérance mathématique de S_n^2 s'écrit alors :

$$E(S_n^2) = \frac{n}{n-1} \sigma^2 - \frac{n}{n-1} \frac{\sigma^2}{n} = \sigma^2 \quad (12)$$

On montre par ailleurs que si X admet un moment d'ordre 4 : $\mu_4 = E[(X - \mu)^4]$, la variance de S_n^{*2} s'écrit¹ :

$$V(S_n^{*2}) = \frac{\mu_4 - \sigma^4}{n} - 2 \frac{(\mu_4 - 2\sigma^4)}{n^2} + \frac{\mu_4 - 3\sigma^4}{n^3} \quad (13)$$

1. Goldberger, Econometrics, John Wiley, 1964, p.97-99.

On peut montrer également que la variance de S_n^2 s'écrit :

$$V(S_n^2) = \frac{\mu_4 - \sigma^4}{n} \quad (14)$$

Si de plus, nous supposons que X est distribué suivant la loi Normale : $X \sim N(\mu, \sigma^2)$, on montre que $\mu_4 = 3\sigma^4$ et $V(S_n^{*2}) = \left(\frac{n-1}{n}\right)^2 \frac{2\sigma^4}{n-1}$ tandis que $V(S_n^2) = \frac{2\sigma^4}{n-1}$.

1.3 Estimateur

Un estimateur $\hat{\theta}$ d'un paramètre inconnu θ est une statistique et donc une variable aléatoire fonction des n variables aléatoires X_1, X_2, \dots, X_n :

$$\hat{\theta}_n = g(X_1, X_2, \dots, X_n) \quad (15)$$

On appelle estimation une réalisation particulière de l'estimateur. Si un autre échantillon aléatoire était tiré sous des conditions identiques, on obtiendrait une réalisation différente de l'échantillon aléatoire et l'estimation serait également différente, présentant des fluctuations d'échantillonnage.

1.4 Distribution d'échantillonnage

On appelle distribution d'échantillonnage la loi de probabilité de la statistique ou de l'estimateur. On peut caractériser cette distribution en calculant analytiquement ses moments, en particulier son espérance mathématique et sa variance ou la spécifier complètement si on connaît la loi des X_i , en utilisant les outils de la statistique mathématique, où si cela est impossible analytiquement, de l'approcher empiriquement par des expériences de simulation qu'on appelle également des expériences de Monte Carlo (cf. infra).

On peut donner une idée intuitive de ce concept de distribution d'échantillonnage de la manière suivante : considérons l'expérience qui consiste à tirer, par exemple, 2000 échantillons aléatoires de taille n sous des conditions identiques. Pour chacun de ces échantillons on peut utiliser un estimateur $\hat{\theta}$ pour calculer une estimation d'un paramètre inconnu θ . Comme les réalisations de ces échantillons diffèrent, ces 2000 estimations seront différentes. La manière dont ces estimations sont distribuées est appelée la distribution d'échantillonnage de l'estimateur $\hat{\theta}$, c'est simplement la fonction de densité de probabilité de $\hat{\theta}$. Elle peut être approchée dans le cadre de notre expérience par la distribution empiriques des fréquences, donc par l'histogramme que l'on peut construire à partir des 2000 estimations de θ .

1.5 Le problème de l'estimation

Le problème de l'estimation est alors celui de la recherche d'une statistique utilisée comme estimateur qui soit dotée de bonnes propriétés en échantillon de taille finie, c'est-à-dire qui conduise à la meilleure évaluation possible, la meilleure estimation, du paramètre inconnu θ . La statistique $\hat{\theta}_n$, estimateur de θ , est une variable aléatoire, dont les réalisations peuvent s'écarter plus ou moins de la vraie valeur du paramètre θ qu'on cherche à estimer. L'objectif pourrait être de choisir un estimateur qui minimise cet écart ou le carré de cet écart, mais il est impossible de définir le meilleur estimateur dans ce sens sauf dans

des cas élémentaires. Ces estimations fluctuent autour de l'espérance mathématique de l'estimateur $E(\hat{\theta}_n)$ avec une dispersion caractérisée par la valeur de la variance $V(\hat{\theta}_n)$ de l'estimateur. Il est donc évident que les estimateurs dont l'espérance mathématique coïncide avec la vraie valeur du paramètre θ et dont la variance est la plus petite possible, sont les plus intéressants.

On établit alors les critères de choix ou propriétés souhaitables pour un estimateur dans un échantillon de taille finie, on aimerait que :

- l'espérance mathématique de l'estimateur soit égale à la vraie valeur du paramètre qu'on cherche à estimer : estimateur centré ou sans biais.
- la variance de l'estimateur soit la plus petite possible : autrement dit la dispersion autour de l'espérance mathématique (la vraie valeur du paramètre) soit la plus petite possible.

Ces critères permettent alors de comparer, en échantillon de taille finie, différents estimateurs possibles du même paramètre inconnu θ .

Parfois, il peut arriver qu'il soit impossible d'exprimer analytiquement les propriétés d'un estimateur dans un échantillon de taille finie, on verra alors qu'on peut comparer les différents estimateurs sur la base de leur propriétés asymptotiques (convergence, efficacité asymptotique).

Exemple 2 Considérons une population caractérisée par une variable aléatoire X (distribuée suivant une loi de probabilité symétrique) d'espérance mathématique μ et de variance σ^2 . On voudrait estimer μ , on dispose d'un échantillon de taille 3 : X_1 , X_2 et X_3 . Considérons les estimateurs suivants :

$$T_1 = \bar{X} = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3 \quad E(T_1) = \mu \quad V(T_1) = \frac{1}{3}\sigma^2 \quad (16)$$

$$T_2 = \frac{1}{6}X_1 + \frac{2}{6}X_2 + \frac{3}{6}X_3 \quad E(T_2) = \mu \quad V(T_2) = \frac{14}{36}\sigma^2 = \frac{1}{3}\sigma^2 + \frac{1}{18}\sigma^2 \quad (17)$$

$$T_3 = X_1 \quad E(T_3) = \mu \quad V(T_3) = \sigma^2 = \frac{1}{3}\sigma^2 + \frac{2}{3}\sigma^2 \quad (18)$$

$$T_4 = M_e \quad E(T_4) = \mu \quad V(T_4) = \frac{\pi^2}{4} \frac{1}{3}\sigma^2 \quad (19)$$

$$T_5 = a \quad a \neq \mu \quad E(T_5) = a \quad V(T_5) = 0 \quad (20)$$

(Pour une distribution symétrique, μ est à la fois la moyenne et la médiane : il peut donc être estimé par la moyenne ou par la médiane de l'échantillon).

Lorsque l'espérance mathématique est égale à la vraie valeur du paramètre, la variance la plus petite donne une plus grande précision autour de la vraie valeur du paramètre. Mais ce critère de variance minimale ne signifie rien tout seul, il n'a de sens qu'utilisé conjointement avec le critère d'absence de biais. Il faut considérer les 2 critères simultanément.

2 Propriétés en échantillon de taille finie

2.1 Estimateur centré - estimateur sans biais

Définition 1 Un estimateur $\hat{\theta}_n$ de θ est dit sans biais ou centré lorsque $E(\hat{\theta}_n - \theta) = 0$ ou encore $E(\hat{\theta}_n) = \theta$.

Ceci signifie que la valeur de l'estimation sur un très grand nombre (une infinité) de réalisations de l'échantillon de taille n fournirait en moyenne la valeur θ (ceci n'est évidemment pas vrai pour une réalisation particulière de l'échantillon et donc de l'estimateur). La propriété d'être sans biais assure que l'estimateur choisi ne conduit pas à une erreur systématique.

Si l'estimateur n'est pas centré, il possède un biais noté : $B(\hat{\theta}_n) = E(\hat{\theta}_n - \theta) = E(\hat{\theta}_n) - \theta$.

Il peut arriver qu'un estimateur soit biaisé, mais que ce biais soit négligeable pour des échantillons de grande taille. Un estimateur est dit asymptotiquement sans biais si : $\lim_{n \rightarrow \infty} B(\hat{\theta}_n) = 0$ soit $\lim_{n \rightarrow \infty} E(\hat{\theta}_n) = \theta$.

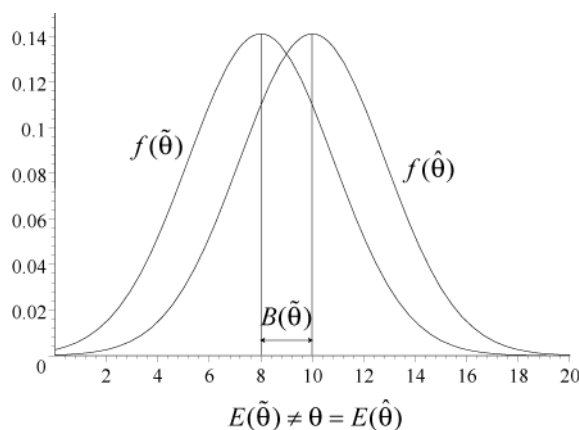


FIGURE 2 – Distribution d'échantillonnage et biais

Exemple 3 La moyenne de l'échantillon \bar{X} est un estimateur centré de la moyenne de la population μ . La variance de l'échantillon S^{*2} est un estimateur biaisé de la variance de la population σ^2 :

$$E(S^{*2}) = \frac{n-1}{n}\sigma^2$$

$$B(S^{*2}) = E(S^{*2}) - \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = -\frac{\sigma^2}{n}$$

Le biais est donc négatif : on sous-estime systématiquement la vraie variance de la population. Cependant si la taille de l'échantillon est très grande, ce biais devient négligeable : S_n^{*2} est donc un estimateur asymptotiquement sans biais de σ^2 . Par contre S_n^2 est un estimateur centré de σ^2 .

2.2 Estimateur efficient ou efficace

Définition 2 Un estimateur $\hat{\theta}_n$ de θ est dit efficient si :

- il est centré.
- parmi tous les estimateurs centrés, il est à variance minimale.

Si de plus on se limite à la classe des estimateurs linéaires, un tel estimateur est appelé *BLUE* (Best Linear Unbiased Estimator).

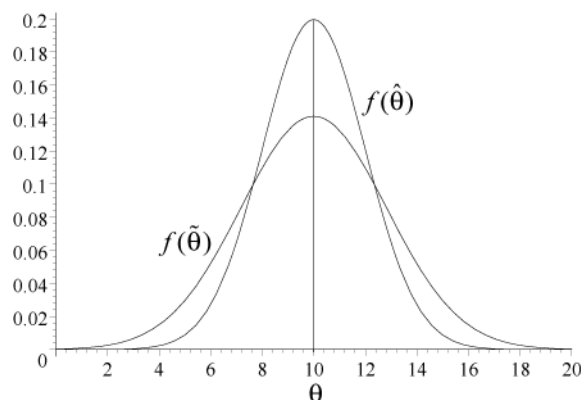


FIGURE 3 – distribution d'échantillonnage et efficacité

Exemple 4 (Dérivation directe d'un estimateur BLUE) Soit une population d'espérance mathématique μ et de variance σ^2 . On voudrait estimer μ . Considérons un échantillon de taille n et définissons un estimateur linéaire de μ .

$$\hat{\mu}_n = \sum_{i=1}^n c_i X_i \quad (21)$$

on a :

$$E(\hat{\mu}_n) = \sum_{i=1}^n c_i E(X_i) = \mu \sum_{i=1}^n c_i \quad (22)$$

$$V(\hat{\mu}_n) = \sum_{i=1}^n c_i^2 V(X_i) = \sigma^2 \sum_{i=1}^n c_i^2 \quad (23)$$

La condition sur les c_i pour que l'espérance mathématique de $\hat{\theta}_n$ soit égale à μ est :

$$E(\hat{\mu}_n) = \mu \quad \Rightarrow \quad \sum_{i=1}^n c_i = 1$$

Parmi tous les estimateurs qui satisfont cette condition, choisissons celui qui a la variance la plus petite : le problème à résoudre s'écrit donc :

$$\begin{cases} \min_{c_i} V(\hat{\mu}_n) = \min_{c_i} \sigma^2 \sum_{i=1}^n c_i^2 = \min_{c_i} \sum_{i=1}^n c_i^2 \\ \text{s.c.} \quad \sum_{i=1}^n c_i = 1 \end{cases} \quad (24)$$

Ecrivons le Lagrangien :

$$\mathbf{L}(c_1, c_2, \dots, c_n; \lambda) = \sum_{i=1}^n c_i^2 + 2\lambda \left[1 - \sum_{i=1}^n c_i \right] \quad (25)$$

Ecrivons maintenant les conditions nécessaires, on annule le gradient du Lagrangien :

$$\begin{cases} \frac{\partial L}{\partial c_i} = 2c_i - 2\lambda = 0 & i = 1, \dots, n \end{cases} \quad (1)$$

$$\begin{cases} \frac{\partial L}{\partial \lambda} = 1 - \sum_{i=1}^n c_i = 0 \end{cases} \quad (2)$$

A partir de (1) on déduit $c_1 = c_2 = \dots = c_n$; ce qui introduit dans (2) donne le point candidat à l'extremum : $c_i^* = \frac{1}{n}$ pour $i = 1, \dots, n$.

Au total, on obtient alors :

$$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}_n \quad (26)$$

La moyenne de l'échantillon \bar{X} est donc bien l'estimateur linéaire centré à variance minimale² de μ : \bar{X} est un estimateur *BLUE* de μ .

2.3 Erreur quadratique moyenne (Mean-Squared Error : MSE)

Nous avons défini un estimateur efficient comme étant celui à variance minimale dans la classe des estimateurs centrés. Mais il est clair qu'on pourrait trouver des estimateurs biaisés à variances plus petites que celle de l'estimateurs centré. Se concentrer uniquement sur les estimateurs centrés pourrait conduire à écarter a priori un estimateur "légèrement" biaisé mais à variance de loin inférieure à celle de l'estimateur centré. Un critère permettant d'arbitrer entre l'absence de biais et la variance est le critère de l'erreur quadratique moyenne.

Définition 3 Soit $\hat{\theta}_n$ un estimateur de θ , l'erreur quadratique moyenne est définie par la quantité : $E(\hat{\theta}_n - \theta)^2$ et mesure la précision de l'estimateur $\hat{\theta}_n$ ou encore le risque d'utiliser $\hat{\theta}_n$ pour estimer θ .

En effet, considérons l'écart entre l'estimateur et la vraie valeur du paramètre :

$$\hat{\theta}_n - \theta = \hat{\theta}_n - E(\hat{\theta}_n) + \underbrace{E(\hat{\theta}_n) - \theta}_{=B(\hat{\theta}_n) \text{ biais}} \quad (27)$$

l'écart quadratique s'écrit alors :

$$(\hat{\theta}_n - \theta)^2 = [\hat{\theta}_n - E(\hat{\theta}_n)]^2 + B^2(\hat{\theta}_n) + 2B(\hat{\theta}_n) [\hat{\theta}_n - E(\hat{\theta}_n)] \quad (28)$$

l'écart quadratique moyen a alors l'expression suivante :

$$E(\hat{\theta}_n - \theta)^2 = E[\hat{\theta}_n - E(\hat{\theta}_n)]^2 + E[B^2(\hat{\theta}_n)] + 2B(\hat{\theta}_n)E[\hat{\theta}_n - E(\hat{\theta}_n)] \quad (29)$$

$$E(\hat{\theta}_n - \theta)^2 = V(\hat{\theta}_n) + B^2(\hat{\theta}_n) + \underbrace{2B(\hat{\theta}_n) [E(\hat{\theta}_n) - E(\hat{\theta}_n)]}_{=0} \quad (30)$$

On peut donc écrire que l'erreur quadratique moyenne, mesure de la précision de l'estimateur, dépend à la fois de la variance et du biais de l'estimateur :

$$E(\hat{\theta}_n - \theta)^2 = V(\hat{\theta}_n) + B^2(\hat{\theta}_n) \quad (31)$$

2. Il s'agit bien d'un minimum global strict parce que le Lagrangien est convexe.

Il n'est pas toujours évident qu'un estimateur sans biais soit préférable à un estimateur biaisé. En effet considérons 2 estimateurs $\hat{\theta}_n$ centré et $\tilde{\theta}_n$ biaisé, mais tels que $V(\tilde{\theta}_n) < V(\hat{\theta}_n)$. On peut alors préférer $\tilde{\theta}_n$ bien qu'il soit biaisé parce que sa variance est plus petite : pour une réalisation de l'échantillon, l'estimation obtenue à partir de $\hat{\theta}_n$ peut être plus éloignée de la vraie valeur du paramètre θ que l'estimation obtenue à partir de $\tilde{\theta}_n$.

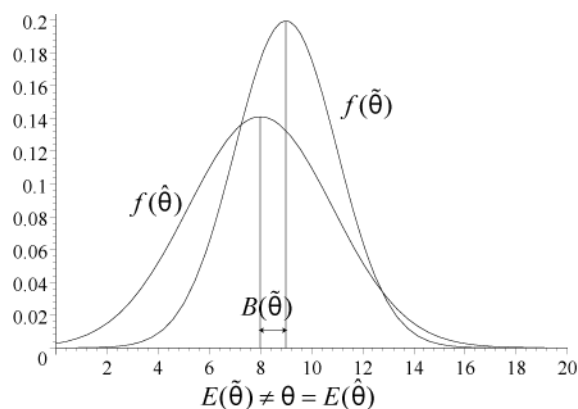


FIGURE 4 – MSE et arbitrage biais - variance

Exemple 5 Soit une population caractérisée par une variable aléatoire X distribuée suivant la Loi Normale $N(\mu, \sigma^2)$. Considérons par exemple les deux estimateurs S^{*2} et S^2 de la variance de la population σ^2 . Nous avons vu que S^{*2} est biaisé alors que S^2 ne l'est pas. Or la variance de S^{*2} est bien inférieure à celle de S^2 :

$$V(S_n^{*2}) = \left(\frac{n-1}{n}\right)^2 \frac{2\sigma^4}{n-1} < V(S_n^2) = \frac{2\sigma^4}{n-1} = MSE(S_n^2) \quad (32)$$

On pourrait comparer les deux estimateurs sur la base du critère de l'erreur quadratique moyenne minimale :

$$MSE(S_n^{*2}) - MSE(S_n^2) = \left[\underbrace{\left(\frac{n-1}{n}\right)^2 \frac{2\sigma^4}{n-1}}_{V(S_n^{*2})} + \underbrace{\frac{\sigma^4}{n^2}}_{B^2(S_n^{*2})} \right] - \frac{2\sigma^4}{n-1} \quad (33)$$

$$= \sigma^4 \left(\frac{2(n-1)+1}{n^2} - \frac{2}{n-1} \right) \quad (34)$$

$$= \sigma^4 \left(\frac{2n-1}{n^2} - \frac{2}{n-1} \right) < 0 \quad (35)$$

L'estimateur biaisé S_n^{*2} est ainsi un peu plus précis. La différence sera évidemment négligeable sur un échantillon de grande taille, mais par exemple le gain relatif de précision est de l'ordre de 10% pour un échantillon de taille 16.

Remarque :

Malheureusement, le critère de l'erreur quadratique moyenne est rarement opérationnel : sauf dans des cas élémentaires, on ne peut pas déterminer le meilleur estimateur satisfaisant ce critère, simplement par qu'il dépend souvent de paramètres inconnus. Aussi est-on moins exigeant en pratique : on utilise le critère d'efficacité et on recherche l'estimateur centré à variance minimale.