Complément de cours: Estimations ponctuelles

Eya ZOUGAR

February 9, 2023

1 Erreur quadratique moyenne EQM

Pour mesurer l'efficacité d'un estimateur par rapport à un autre, on peut utiliser l'erreur quadratique moyenne.

Definition 1 Soit T_n un estimateur du paramètre θ. Son erreur quadratique moyenne est définie par:

$$EQM = \mathbb{E}\left[(T_n - \theta)^2\right]$$

On dira que T_n converge en moyenne quadratique vers θ lorsque son erreur quadratique moyenne convergera vers 0.

Décomposition Biais-Variance de l'erreur quadratique moyenne:

$$\mathbb{E}\left[(T_n - \theta)^2\right] = \underbrace{\mathbb{E}\left[(T_n - \mathbb{E}(T_n))^2\right]}_{\text{Variance}} + \underbrace{(\mathbb{E}[T_n] - \theta)^2}_{\text{Biais}}.$$

Preuve: On va utiliser le faite que $\mathbb{E}[a] = a$ et $\mathbb{E}[aX + Y] = a\mathbb{E}[X] + \mathbb{E}[Y]$

$$\begin{aligned}
\mathbb{E}\left[(T_n - \theta)^2\right] &= \mathbb{E}\left[(T_n - \mathbb{E}(T_n) + \mathbb{E}(T_n) - \theta)^2\right] \\
&= \mathbb{E}\left[(T_n - \mathbb{E}(T_n))^2\right] + \mathbb{E}\left[(\mathbb{E}[T_n] - \theta)^2\right] + 2\mathbb{E}\left[(T_n - \mathbb{E}(T_n))(\mathbb{E}[T_n] - \theta)\right] \\
&= \mathbb{E}\left[(T_n - \mathbb{E}(T_n))^2\right] + (\mathbb{E}[T_n] - \theta)^2 + 2\mathbb{E}\left[(T_n - \mathbb{E}(T_n))(\mathbb{E}[T_n] - \theta)\right] \\
&= \mathbb{E}\left[(T_n - \mathbb{E}(T_n))^2\right] + (\mathbb{E}[T_n] - \theta)^2 + \underbrace{2(\mathbb{E}[T_n] - \theta)\,\mathbb{E}[T_n - \mathbb{E}(T_n)]}_{0} \\
&= \mathbb{E}\left[(T_n - \mathbb{E}(T_n))^2\right] + (\mathbb{E}[T_n] - \theta)^2
\end{aligned}$$

Si l'estimateur est sans biais ($\mathbb{E}[T_n] = \theta$), cette égalité se réduit à:

$$\mathbb{E}\left[(T_n - \theta)^2\right] = \mathbb{V}\text{ar}(T_n).$$

2 L'estimateur de la moyenne.

Soient $X_1, X_2, \dots, X_n$ des variables aléatoires réelles, indépendantes et de même loi, d'espérance μ et de variance σ^2. On souhaite estimer l'espérance μ, supposée inconnue.

Pour estimer le paramètre $\mu = \mathbb{E}[X_1]$, on utilise la méthode des moments avec $h = Id$.

Definition 2 Estimation de la moyenne Pour estimer l'espérance μ, on prend la moyenne empirique des (X_j), notée $\overline{X}_n$ et définie par :

$$\overline{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$$

1. **L'estimateur $\overline{X}_n$ est un estimateur sans biais du paramètre μ.**
   En effet, on a puisque les (X_i) sont i.i.d

$$\mathbb{E}(\overline{X}_n) = \mathbb{E}\left[\frac{1}{n}\sum_{j=1}^{n} X_j\right] = \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}(X_j) = \frac{1}{n}\sum_{j=1}^{n}\mu = \mu$$

2. **Convergence en moyenne qudratique:**
   Sa variance tend vers 0 lorsque n tend vers l'infini.
   En effet:

$$\mathbb{V}\mathrm{ar}(\overline{X}_n) = \mathbb{V}\mathrm{ar}\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right) = \frac{1}{n^2}\sum_{j=1}^{n}\mathbb{V}\mathrm{ar}(X_j) = \frac{\sigma^2}{n}$$

   D'aprés (1) et (2) on déduit que l'estimateur $\overline{X}_n$ converge en moyenne quadratique vers μ.

3. **Convergence en Loi:**
   Puisque les (X_i) sont i.i.d. et $\mathbb{E}(\overline{X}_1) < \infty$ alors les hypothèses du TLC son satisfaites. Donc, d'aprés TLC, on aura:

$$\sqrt{n}\left(\overline{X}_n - \mu\right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2)$$

4. D'aprés l'inégalité de tchybechev, on aura $\forall \epsilon > 0$

$$\mathbb{P}\left(|\overline{X}_n - \mathbb{E}(\overline{X}_n)| > \epsilon\right) \leq \frac{\mathbb{V}\mathrm{ar}(\overline{X}_n)}{\epsilon^2} \Longleftrightarrow \mathbb{P}\left(|\overline{X}_n - \mu| > \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2} \xrightarrow[n\to\infty]{} 0$$

   ce qui montre que $\overline{X}_n$ converge en probabilité vers μ.

5. **Convergence en p.s:**
   Une application directe du théorème des grands nombres, montre: $\overline{X}_n \xrightarrow[n\to\infty]{p.s.} \mu$.

3 Estimation de la variance

Soient $X_1, X_2, \ldots, X_n$ des variables aléatoires réelles, indépendantes et de même loi, d'espérance μ et de variance σ^2.

On suppose que μ et σ^2 sont inconnus.

On a $\sigma^2 = \mathbb{V}\mathrm{ar}(X_1) = \mathbb{E}\left((X_1 - \mu)^2\right)$. On souhaite estimer la variance σ^2.

Definition 3
- Supposons ds un premier temps μ connu. Estimer σ^2, c'est en fait estimer une espérance, et donc un estimateur naturel est:

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j - \mu)^2$$

- Cependant, en génèral μ est inconnu. On l'estime par $\overline{X}_n$ et un estimateur de la variance σ^2 est alors donné par :

$$\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X}_n)^2 = \frac{1}{n}\sum_{j=1}^{n}X_j^2 \ - \ \overline{X}_n^2.$$

  Cette méthode, qui consiste à remplacer un paramètre par son estimateur, est dite "du plug-in".

Dans la suite on va utiliser **la décomposition suivante:**

$$\sum_{j=1}^{n}(X_j - \overline{X}_n)^2 \ = \ \sum_{j=1}^{n}(X_j - \mu)^2 \ - \ n(\overline{X}_n - \mu)^2 \tag{1}$$

En effet:
$$\sum_{j=1}^{n}(X_j-\overline{X}_n)^2=\sum_{j=1}^{n}(X_j-\mu+\mu-\overline{X}_n)^2=\sum_{j=1}^{n}(X_j-\mu)^2+\sum_{j=1}^{n}(\mu-\overline{X}_n)^2+2\sum_{j=1}^{n}(X_j-\mu)(\mu-\overline{X}_n)$$
Or on a $\sum_{j=1}^{n}(\mu-\overline{X}_n)^2=(\mu-\overline{X}_n)^2\sum_{j=1}^{n}1=n\ (\mu-\overline{X}_n)^2$ et
$$\begin{aligned}
\sum_{j=1}^{n}(X_j-\mu)(\mu-\overline{X}_n) &= (\mu-\overline{X}_n)\sum_{j=1}^{n}(X_j-\mu)\\
&= (\mu-\overline{X}_n)(\sum_{j=1}^{n}X_j-\mu\sum_{j=1}^{n}1)\\
&= (\mu-\overline{X}_n)(n\overline{X}_n-n\mu)\\
&= -2n(\overline{X}_n-\mu)^2
\end{aligned}$$

1. **Cet estimateur est biaisé, mais asymptotiquement sans biais.**

   En effet, à partir de la décomposition (1), on aura:
$$\widehat{\sigma}^2=\frac{1}{n}\sum_{j=1}^{n}(X_j-\overline{X}_n)^2\ =\ \frac{1}{n}\sum_{j=1}^{n}(X_j-\mu)^2\ -\ (\overline{X}_n-\mu)^2$$
   alors,
$$\begin{aligned}
\mathbb{E}\left(\widehat{\sigma}^2\right) &= \frac{1}{n}\mathbb{E}\left(\sum_{j=1}^{n}(X_j-\mu)^2\right)+\mathbb{E}\left((\overline{X}_n-\mu)^2\right)\\
&= \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left((X_j-\mu)^2\right)+\mathbb{V}\mathrm{ar}(\overline{X}_n)\\
&= \frac{1}{n}\sum_{j=1}^{n}\mathbb{V}\mathrm{ar}(X_1)+\mathbb{V}\mathrm{ar}(\overline{X}_n)\ \text{car les }\ X_i\text{ sont i.i.d}\\
&= \frac{1}{n}n\sigma^2+\frac{\sigma^2}{n}\\
&= \frac{n-1}{n}\ \sigma^2\ \underset{n\to\infty}{\longrightarrow}\ \sigma^2
\end{aligned}$$

2. **Calcul de la variance:** On note $\tau^4=\mathbb{E}\left[(X_1-\mu)^4\right]$.
   On calcule d'abord:
   $\mathbb{V}\mathrm{ar}\left((X_1-\mu)^2\right)=\mathbb{E}\left((X_1-\mu)^4\right)-\left(\mathbb{E}\left((X_1-\mu)^2\right)\right)^2=\tau^4-\mathbb{V}\mathrm{ar}(X_1)^2=\tau^4-\sigma^4.$

   On a, à partir de la décomposition (1),
$$\widehat{\sigma}^2=\frac{1}{n}\sum_{j=1}^{n}(X_j-\overline{X}_n)^2\ =\ \frac{1}{n}\sum_{j=1}^{n}(X_j-\mu)^2\ -\ (\overline{X}_n-\mu)^2$$
   avec $\mathbb{V}\mathrm{ar}(aX+bY)=a^2\mathbb{V}\mathrm{ar}(X)+2ab\ \mathbb{C}\mathrm{ov}(X,Y)+b^2\mathbb{V}\mathrm{ar}(X)$
$$\begin{aligned}
\mathbb{V}\mathrm{ar}(\widehat{\sigma}^2) &= \mathbb{V}\mathrm{ar}\left(\frac{1}{n}\sum_{j=1}^{n}(X_j-\mu)^2\right)+\mathbb{V}\mathrm{ar}\left((\overline{X}_n-\mu)^2\right)\\
&- \frac{2}{n}\mathbb{C}\mathrm{ov}\left(\textstyle\sum_{j=1}^{n}(X_j-\mu)^2\,;\,(\overline{X}_n-\mu)^2\right)\\
&= A+B-2C
\end{aligned}$$
   Puisque les variables (X_j) sont i.i.d, on a: $A=\frac{1}{n}\mathbb{V}\mathrm{ar}\left((X_1-\mu)^2\right)=\frac{\tau^4-\sigma^4}{n}$

$$\begin{aligned} C &= \frac{1}{n}\sum_{j=1}^{n}\mathbb{C}\mathrm{ov}\left((X_j-\mu)^2\,;\,(\overline{X}_n-\mu)^2\right) \\ &= \mathbb{C}\mathrm{ov}\left((X_1-\mu)^2\,;\,(\overline{X}_n-\mu)^2\right) \\ &= \frac{1}{n^2}\,\mathbb{C}\mathrm{ov}\left((X_1-\mu)^2\,;\,\left(\sum_{j=1}^{n}(X_j-\mu)\right)^2\right) \end{aligned}$$

Or

$$\left(\sum_{j=1}^{n}(X_j-\mu)\right)^2 = \sum_{j=1}^{n}(X_j-\mu)^2 + 2\sum_{j<k}(X_j-\mu)(X_k-\mu)$$

Comme les variables (X_j) sont indépendantes, on a

$$\mathbb{C}\mathrm{ov}\left((X_1-\mu)^2\,;\,(X_j-\mu)^2\right) = 0$$

et pour tout $1\le j<k\le n$, $\mathbb{C}\mathrm{ov}\left((X_1-\mu)^2\,;\,(X_j-\mu)(X_k-\mu)\right)=0$.
Ainsi, en rassemblant les différents résultats, on obtient que

$$C = \frac{\mathbb{V}\mathrm{ar}((X_1-\mu)^2)}{n^2} = \frac{\tau^4-\sigma^4}{n^2}$$

Il reste maintenant à calculer $B=\mathbb{V}\mathrm{ar}\left((\overline{X}_n-\mu)^2\right)$. On a

$$\begin{aligned} B &= \mathbb{C}\mathrm{ov}\left((\overline{X}_n-\mu)^2\,;\,(\overline{X}_n-\mu)^2\right) \\ &= \frac{1}{n^4}\mathbb{C}\mathrm{ov}\left(\left(\sum_{j=1}^{n}(X_j-\mu)\right)^2\,;\,\left(\sum_{j=1}^{n}(X_j-\mu)\right)^2\right) \end{aligned}$$

en développant et en utilisant le fait que les (X_j) sont indépendantes, on montre facilement que

$$B = \frac{\tau^4+2\sigma^4}{n^3}$$

3. **Convergence presque sûre de $\widehat{\sigma}^2$:**
   On a montré d'aprés section 2, que $\overline{X}_n \xrightarrow[n\to\infty]{p.s.} \mu$.
   En appliquant la loi forte des grands nombres, pour la suite de variables aléatoires $\left((X_n-\mu)^2\right)_{n\ge1}$, on aura:
   $$\frac{1}{n}\sum_{j=1}^{n}(X_j-\mu)^2 \xrightarrow[n\to\infty]{p.s.} \mathbb{E}\left[(X_1-\mu)^2\right]=\sigma^2$$
   D'où, d'aprés la décomposition (1),
   $$\widehat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j-\overline{X}_n)^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j-\mu)^2 - (\overline{X}_n-\mu)^2 \xrightarrow[n\to\infty]{p.s.} \sigma^2$$

4. **Convergence en probabilité de la suite $\sqrt{n}(\overline{X}_n-\mu)^2$:** $\sqrt{n}(\overline{X}_n-\mu)^2 \xrightarrow[n\to\infty]{P} 0$. En effet:
   D'aprés l'inégalité de Markov, on a $\forall\epsilon>0$
   $$\mathbb{P}\left(|\sqrt{n}(\overline{X}_n-\mu)^2|>\epsilon\right) \le \frac{\sqrt{n}\;\mathbb{E}[(\overline{X}_n-\mu)^2]}{\epsilon} = \frac{\sqrt{n}\mathbb{V}\mathrm{ar}(\overline{X}_n)}{\epsilon} = \frac{\sigma^2}{\sqrt{n}\epsilon^2} \xrightarrow[n\to\infty]{} 0$$

5. **Convergence en Loi:**
   En appliquant le théorème de limite centrale, pour la suite de variables aléatoires $\left((X_n-\mu)^2\right)_{n\ge1}$, on aura:

$$\sqrt{n}\left(\frac{1}{n}\sum_{j=1}^{n}(X_j-\mu)^2 \;-\; \sigma^2\right) \;\xrightarrow[n\to\infty]{\mathcal{L}}\; \mathcal{N}(0,\tau^4-\sigma^4),$$

D'aprés, (4), on peut déduire que: $\sqrt{n}(\overline{X}_n-\mu)^2 \xrightarrow[n\to\infty]{\mathcal{L}} 0$.
D'où, d'aprés la décomposition (1),

$$\sqrt{n}\left(\widehat{\sigma}^2-\sigma^2\right)=\sqrt{n}\left(\frac{1}{n}\sum_{j=1}^{n}(X_j-\overline{X}_n)^2-\sigma^2\right)=\sqrt{n}\left(\frac{1}{n}\sum_{j=1}^{n}(X_j-\mu)^2 \;-\sigma^2-(\overline{X}_n-\mu)^2\right) \;\xrightarrow[n\to\infty]{\mathcal{L}}\; \mathcal{N}(0,\tau^4-\sigma^4)$$