(*) $S_n^2 := \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ est un

estimateur biaisé de σ^2, variance commune

des $(X_i)_{i=1...n}$

<u>démo:</u> $E[S_n^2] = \frac{1}{n}\sum_{i=1}^{n} E[(X_i - \overline{X})^2]$

$= \frac{1}{n}\sum_{i=1}^{n} E[(X_i - m)^2 + (m - \overline{X})^2 + 2(X_i - m)(m - \overline{X})]$

où $m = E[X_i]$ est la moyenne théorique commune des $(X_i)_{i=1...n}$

<u>Ainsi:</u>

$$E[S_n^2] = \frac{1}{n}\sum_{i=1}^{n}\{Var(X_i - m) + Var(m - \overline{X}) - 2E[(X_i - m)(\overline{X} - m)]\}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\{Var(X_i) + Var(\overline{X}) - \frac{2}{n}\sum_{j=1}^{n} E[(X_i - m)(X_j - m)]\}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\{\sigma^2 + \frac{\sigma^2}{n} - \frac{2}{n}E[(X_i - m)^2]\}$$

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

$$= \frac{n-1}{n}\sigma^2$$

remarquez alors que S_{n-1}^2 est un estimateur non biaisé de σ^2 !

Durée d d'un feu rouge à estimer

↳ ① On a vu que le temps d'attente des i^{e} automobiliste était bien modélisé par $T_i \sim Unif([0;d])$ de densité $f_{T_i}(t) = \frac{1}{d} \mathbb{1}_{[0;d]}(t)$.

On calcule aisément $E(T_i) = \frac{d}{2}$ et $var(T_i) = \frac{d^2}{12}$.

② On voit simplement que $E[\overline{T}] = \frac{d}{2}$ et $var(\overline{T}) = \frac{d^2}{12n}$ (modulo $var(\sum_i T_i) = \sum_i var(T_i)$ car les T_i sont indépendants)

$\hat{\theta}_1 := 2 \times \overline{T}$ est un estimateur non biaisé de d

Comme $Var(\hat{\theta}_1) = \frac{d^2}{3n} \xrightarrow{n \to +\infty} 0$, $\hat{\theta}_1$ converge en proba vers d. (utiliser l'inégalité de Bienaymé-Tchebichev)

③ • $M_n := \sup_{i=1 \dots n} (T_i) \leq t \iff T_i \leq t, \forall i = 1 \dots n$

Comme les $(T_i)_i$ sont indépendants : $P(M_n \leq t) = \prod_{i=1}^{n} P(T_i \leq t)$

↳ $P(M_n \leq t) = \begin{cases} 0, & \text{si } t < 0 \\ \left(\frac{t}{d}\right)^n, & \text{si } t \in [0;d] \\ 1, & \text{si } t > d \end{cases}$

En dérivant, la densité de M_n est $f_{M_n}(u) = \frac{n \cdot u^{n-1}}{d^n} \mathbb{1}_{[0;d]}$

• Le graphe de $f_{M_{30}}$ est "plus pointu" que celui de f_{M_3} ; la variance en $\frac{1}{n^2}$ ci-après (écart-type en $\frac{1}{n}$) indique que M_{30} est à peu près 10 fois plus concentré au voisinage de d que M_3 ($\frac{n+1}{n} \cdot M_n$ en fait)

$$E[M_n] = \frac{n}{d^n}\int_0^d u \cdot u^{n-1}\,du = \frac{n \times d}{n+1}$$

$$\mathrm{var}(M_n) = \underbrace{E[M_n^2]}_{} - E[M_n]^2$$

$$= \frac{n}{d^n}\int_0^d u^2 u^{n-1}\,du = \frac{n\, d^2}{n+2}$$

$$\hookrightarrow \mathrm{var}(M_n) = \frac{n\, d^2}{(n+1)^2(n+2)}.$$

On propose $\hat{\theta}_2 := \frac{n+1}{n} M_n$ comme estimateur non biaisé de d, de variance $\mathrm{var}(\hat{\theta}_2) = \frac{d^2}{n(n+2)} \xrightarrow{n\to+\infty} 0$ qui converge bien en probabilité vers d.

(4) $\mathrm{var}(\hat{\theta}_1) - \mathrm{var}(\hat{\theta}_2) = d^2\left(\frac{1}{3n} - \frac{1}{n(n+2)}\right) \leq 0$

avec égalité ssi $n=1$.

$\to$ On préfèrera toujours $\hat{\theta}_2$.

(5) Appli numérique :

$\hat{\theta}_1 = 22$ n'est vraiment pas terrible puisqu'on a $t_2 = 24$ (et donc $d \geq 24$) !!

En revanche, $\hat{\theta}_2 = \frac{4}{3}\times 24 = 32$. Pq pas...

(6) Je m'aperçois que j'oubliais pas vous demander si $\hat{\theta}_2$ était efficace... La réponse est non car $V_n(d)$ serait en $\frac{1}{n^3}$ $\left(V_n(d) = \frac{1}{-nE\left[\frac{\partial^2 \log f_{M_n}(u)}{}\right]}\right)$

Estimation de l'écart-type d'une loi $|\mathcal{N}|$

$X_1, \dots, X_n$ des variables aléatoires indépendantes et identiquement distribuées (iid) $\mathcal{N}(m, \sigma^2)$. On observe $\forall i$ les $Y_i = |X_i - m|$.

- Si X de densité f_X, $|X|$ a une densité double sur $\mathbb{R}^+$, nulle sur $\mathbb{R}^-$ $\to$ $f_{|X|} = 2 f_X \cdot \mathbb{1}_{\mathbb{R}^+}$

Or $\forall i$, $X_i - m \sim \mathcal{N}(0 ; \sigma^2)$

$\hookrightarrow \forall i$, $d_{Y_i}(y) = \sqrt{\dfrac{2}{\pi \sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\, \mathbb{1}_{\mathbb{R}^+}(y)$

$E[Y_i] = \int_{\mathbb{R}^+} \sqrt{\dfrac{2}{\pi\sigma^2}}\, y\, e^{-y^2/2\sigma^2}\, dy$; on pose $u = \dfrac{y^2}{2\sigma^2}$

$\left(du = \dfrac{y\,dy}{\sigma^2}\right)$

$= \sqrt{\dfrac{\sigma^2}{2\pi}} \times 2 \underbrace{\int_{\mathbb{R}^+} e^{-u}\, du}_{= 1}$

$\Rightarrow \forall i$, $E[Y_i] = \sigma \sqrt{\dfrac{2}{\pi}}$

- Si $\hat{\sigma} = \sum_{i=1}^n a_i Y_i$ comme moyenne pondérée des Y_i pour estimer σ, on a que $E[\hat{\sigma}] = E[Y_i] \cdot \left(\sum_{i=1}^n a_i\right)$

Donc $\hat{\sigma}$ sans biais $\Leftrightarrow \sum_{i=1}^n a_i = \sqrt{\dfrac{\pi}{2}}$

- $a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$. Minimiser $\sum_{i=1}^n a_i^2$ sous la contrainte $\sum_{i=1}^n a_i = \frac{\pi}{2}$ équivaut à minimiser la distance de l'origine à l'hyperplan affine H $\sum_{i=1}^n a_i = \sqrt{\frac{\pi}{2}}$, convexe complet.

On $\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix} \perp H$. Par théorème du projeté, O se projetté sur H en un point de coordonnées proportionnelles à $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$. Donc $\boxed{a_i = \text{constante} = \frac{1}{n} \cdot \sqrt{\frac{\pi}{2}} \quad \forall i}$

<u>rque:</u> on aurait aussi pu utiliser le lagrangien du problème...

- $Var\left(\hat{\sigma}\right) = \frac{1}{n^2} \frac{\pi}{2} \times n \cdot var(Y_i)$ alors, puisque les $(Y_i)_{i=1 \dots n}$ sont indépendantes.

$$var(Y_i) = \underbrace{E[Y_i^2]}_{=\sigma^2} - \underbrace{E[Y_i]}_{=\sigma\sqrt{\frac{2}{\pi}}}{}^2 = \sigma^2 \cdot \frac{\pi - 2}{\pi}$$

$$\Rightarrow Var\left(\hat{\sigma}\right) = \frac{\pi - 2}{2n} \cdot \sigma^2$$

- Les hypothèses (H_1) à (H_5) du théorème de Cramer-Rao sont satisfaites et

$$\frac{\partial \ln\left[d_{Y_i}(y)\right]}{\partial \sigma} = -\frac{1}{\sigma} + \frac{y^2}{\sigma^3}$$

$$\frac{\partial^2 \ln\left[d_{Y_i}(y)\right]}{\partial \sigma^2} = \frac{1}{\sigma^2} - \frac{3y^2}{\sigma^4}$$

$$\Rightarrow E\left[\frac{\partial^2 \ln\left[d_{Y_i}(y)\right]}{\partial \sigma^2}\right] = \frac{1}{\sigma^2} - \frac{3}{\sigma^2} = -\frac{2}{\sigma^2}$$

car $E[Y_i^2] = \sigma^2$

$$\Rightarrow V_n(\sigma) = \frac{1}{-n E\left[\frac{\partial^2 \ln d_{Y_i}}{\partial \sigma^2}\right]} = \frac{\sigma^2}{2n}$$

Comme $\frac{Var\, \hat{\sigma}}{V_n(\sigma)} = \pi - 2 \approx 1,14 > 1$, $\hat{\sigma}$ n'est pas efficace (à une constante près).

III Estimateur du Maximum de Vraisemblance (EMV, MLE en anglais) d'une proportion

On veut estimer la proportion p de pièces défectueuses à la sortie d'une chaîne d'usine. On prélève n_1 pièces. On modélise l'état de la $i^{\text{ème}}$ pièce tirée par $X_i = \begin{cases} 1, & \text{si la pièce est défectueuse} \\ 0, & \text{si elle est OK.} \end{cases}$

$X_i = p.\delta_1 + (1-p).\delta_0$. Comme les pièces sont indépendantes, la vraisemblance vaut

$$L(x;p) = p^{\overbrace{x_1+\dots+x_n}^{\text{défectueuses}}} (1-p)^{\overbrace{n-x_1-\dots-x_n}^{\text{OK}}}$$

$$\rightarrow \frac{\partial \ln L}{\partial p} = \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1-p} \quad \text{qui}$$

s'annule pour $(1-p)\sum_{i=1}^n x_i = p\left[n - \sum_{i=1}^n x_i\right]$

$$\Rightarrow \boxed{\hat{p} = \frac{\sum_{i=1}^n x_i}{n}} \quad (\text{o-} \int !)$$

Propriétés de cet estimateur : $E[\hat{p}] = p$ (sans biais)

$$\text{var}[\hat{p}] = \frac{p(1-p)}{n}$$

Enfin $E\left[\frac{\partial^2 \ln L}{\partial p^2}\right] = -\frac{n}{p(1-p)} \rightarrow V_n(p) = \frac{p(1-p)}{n} = \text{var}(\hat{p})$

$\rightarrow$ $\hat{p}$ est efficace (le cours ne nous le donnait qu'asymptotiquement efficace)