Propriétés de l'estimateur de la moyenne d'une population quantitative univariée

ECHANTILLON ALEATOIRE SIMPLE (EAS)

① La moyenne d'échantillonnage de la moyenne d'échantillon est égale à la moyenne de la population

$$\boxed{E_{EAS}(\hat{\mu}) = E_{EAS}(\overline{X}) = \mu}$$

Démonstration :

Par hypothèse : échantillon ≡ valeurs réalisées de n v.a. iid $(X_1, X_2, \ldots, X_n)$ telles que $E(X_i) = \mu$ et $var(X_i) = \sigma^2, \forall i$

$$E(\overline{X}) = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] \underset{E \text{ linéaire}}{=} \frac{1}{n}\sum_{i=1}^{n} E(X_i) \underset{hyp}{=} \frac{1}{\cancel{n}}\cancel{n}\mu = \mu$$

② La variance d'échantillonnage de la moyenne d'échantillon est égale à la variance de la population divisée par la taille de l'échantillon.

$$\boxed{\underset{EAS}{Var}(\hat{\mu}) = \underset{EAS}{var}(\overline{X}) = \frac{\sigma^2}{n}}$$

Démonstration:

$$var(\overline{X}) = var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

$var(aX) = a^2 \, var\, X$

$$= \frac{1}{n^2} var\left(\sum_{i=1}^{n} X_i\right)$$

$var(X+Y) = var\,X + var\,Y + 2\,cov(X,Y)$

mais $cov(X,Y) = 0$ lorsque les v.a. sont indp$^{\underline{dtes}}$ (cf hyp)

$$= \frac{1}{n^2}\sum_{i=1}^{n} var\, X_i$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2$$

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

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

ECHANTILLON ALEATOIRE SANS REMISE (EASR)

① La moyenne d'échantillonnage de la moyenne d'échantillon est égale à la moyenne de la population

$$\boxed{E_{EASR}(\bar{X}) = \mu}$$

Dém : $E(\bar{X}) = E\left(\frac{1}{n}\sum_{j=1}^{N} \underline{n}_j x_j\right)$

$= \frac{1}{n}\sum_{j=1}^{N} \underbrace{E(\underline{n}_j)}_{0 \cdot (1-\frac{n}{N}) + 1\cdot\frac{n}{N}} x_j$ $\qquad \underline{n}_j \sim \begin{pmatrix} 0 & 1 \\ 1-\frac{n}{N} & \frac{n}{N} \end{pmatrix}$

$= \frac{1}{\cancel{n}} \frac{\cancel{n}}{N} \sum_{j=1}^{N} x_j$

$= \mu$

② La variance d'échantillonnage de la moyenne d'échantillon est égale au produit de 2 facteurs :

- l'un est la variance de la population divisée par la taille de l'échant.
- l'autre dépend de la taille de la population et de la taille de l'échant.

et est tel que si $N \gg n \Rightarrow var_{EASR}(\underline{\bar{X}}) \simeq var_{EAS}(\underline{\bar{X}})$

Remarque préliminaire : $\left(\sum_{j=1}^{N} \underline{n}_j x_j\right)^2 = ??$

Ds le cas particulier où $N = 3$, on aurait :

$\left(\sum_{j=1}^{3} \underline{n}_j x_j\right)^2 = \underline{n}_1^2 x_1^2 + \underline{n}_2^2 x_2^2 + \underline{n}_3^2 x_3^2 + 2\underline{n}_1\underline{n}_2 x_1 x_2 + 2\underline{n}_2\underline{n}_3 x_2 x_3 + 2\underline{n}_1\underline{n}_3 x_1 x_3$

$= \sum_{j=1}^{3} \underline{n}_j^2 x_j^2 \quad \underbrace{\begin{aligned} &+ \underline{n}_1 x_1 \underline{n}_2 x_2 + \underline{n}_2 x_2 \underline{n}_1 x_1 \\ &+ \underline{n}_2 \underline{n}_3 x_2 x_3 + \underline{n}_3 x_3 \underline{n}_2 x_2 \\ &+ \underline{n}_1 x_1 \underline{n}_3 x_3 + \underline{n}_3 x_3 \underline{n}_1 x_1 \end{aligned}}_{\text{tous les couples possibles construits avec } k \text{ et } j \text{ en imposant que les indices } \neq}$

$= \sum_{j=1}^{3} \underline{n}_j^2 x_j^2 + \sum_{j=1}^{3}\sum_{k=1, k\neq j}^{3} \underline{n}_j \underline{n}_k x_j x_k$

$$\boxed{\mathrm{var}_{EASR}(\bar{X}) = \frac{\sigma^2}{n}\,\frac{N-n}{N-1}}$$

<u>Dém</u> : $\mathrm{var}(\bar{X}) = E(\bar{X}^2) - [E(\bar{X})]^2$ (**)

Or
$$E(\bar{X}^2) = E\left[\left(\frac{1}{n}\sum_{j=1}^{N} \underline{n}_j\, x_j\right)^2\right]$$

$$= \frac{1}{n^2}\, E\left(\sum_{j=1}^{N} \underline{n}_j\, x_j\right)^2$$

cf REMARQUE PRÉLIMINAIRE

$$= \frac{1}{n^2}\, E\left(\sum_{j=1}^{N} \underline{n}_j^2 x_j^2 + \sum_{j=1}^{N}\sum_{\substack{k=1\\ j\neq k}}^{N} \underline{n}_j\,\underline{n}_k\, x_j\, x_k\right)$$

$$= \frac{1}{n^2}\left[\sum_{j=1}^{N} E(\underline{n}_j^2 x_j^2) + \sum_{j=1}^{N}\sum_{k=1}^{N} E(\underline{n}_j\,\underline{n}_k\, x_j\, x_k)\right]$$

$$E(\underline{n}_j^2 x_j^2) = x_j^2\, E(\underline{n}_j^2) = x_j^2\, E(\underline{n}_j) = x_j^2\,\frac{n}{N}$$

$$\underline{n}_j^2 = \underline{n}_j$$

$$E(\underline{n}_j\,\underline{n}_k\, x_j\, x_k) = x_j\, x_k\, E(\underline{n}_j\,\underline{n}_k)_{j\neq k} = x_j\, x_k\, \frac{n(n-1)}{N(N-1)}$$

$$\underline{n}_j \sim \begin{pmatrix} 0 & 1 \\ 1-\frac{n}{N} & \frac{n}{N}\end{pmatrix}$$

$$\rightarrow E(\underline{n}_j) = 0.\left(1-\frac{n}{N}\right) + 1.\frac{n}{N} = \frac{n}{N}$$

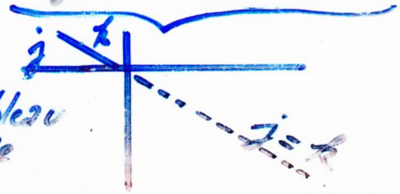
$$\underline{n}_j\,\underline{n}_k \sim \begin{pmatrix} 0 & 1 \\ 1-\frac{n(n-1)}{N(N-1)} & \frac{n(n-1)}{N(N-1)}\end{pmatrix}$$

$$\rightarrow E(\underline{n}_j\,\underline{n}_k) = \frac{n(n-1)}{N(N-1)}$$

$$= \frac{1}{n^2}\left[\sum_{j=1}^{N} \frac{n}{N}\, x_j^2 + \sum_{j=1}^{N}\sum_{\substack{k=1\\ j\neq k}}^{N} \frac{n(n-1)}{N(N-1)}\, x_j\, x_k\right]$$

$$= \frac{1}{n^2}\left[\sum_{j=1}^{N} \frac{n}{N}\, x_j^2 + \frac{n(n-1)}{N(N-1)} \underbrace{\sum_{j=1}^{N}\sum_{\substack{k=1\\ j\neq k}}^{N} x_j\, x_k}_{(*)}\right]$$

(*) il faut faire la somme sur t^t le tableau <u>SAUF</u> sur la diagonale



$$(*) = \sum_{j=1}^{N} \sum_{k=1}^{N} x_j \cdot x_k - \sum_{j=1}^{N} x_j^2$$

$$= \frac{1}{n^2}\left[\sum_{j=1}^{N} \frac{n}{N} x_j^2 + \frac{n(n-1)}{N(N-1)}\left(\sum_{j=1}^{N}\sum_{k=1}^{N} x_j \cdot x_k - \sum_{j=1}^{N} x_j^2\right)\right]$$

$$= \frac{1}{n^{\cancel{2}}} \frac{\cancel{n}}{N(N-1)}\left[\sum_{j=1}^{N} (N-1)\, x_j^2 + (n-1)\left(\sum_{j=1}^{N}\sum_{k=1}^{N} x_j \cdot x_k - \sum_{j=1}^{N} x_j^2\right)\right]$$

$$= \frac{1}{nN(N-1)}\left[\underbrace{\big((N-1)-(n-1)\big)}_{(N-n)}\sum_{j=1}^{N} x_j^2 + (n-1)\sum_{j=1}^{N}\sum_{k=1}^{N} x_j \cdot x_k\right]$$

$$= \frac{1}{nN(N-1)}\left[(N-n)\sum_{j=1}^{N} x_j^2 + (n-1)\underbrace{\sum_{j=1}^{N} x_j}_{(x_1+\dots+x_N)}\ \underbrace{\sum_{k=1}^{N} x_k}_{(x_1+\dots+x_N)}\right]$$

$$(x_1+\dots+x_N)\cdot(x_1+\dots+x_N)$$
$$= (x_1+\dots+x_N)^2$$
$$= \left(\sum_{j=1}^{N} x_j\right)^2$$

$$= \frac{1}{nN(N-1)}\left[(N-n)\sum_{j=1}^{N} x_j^2 + (n-1)\left(\sum_{j=1}^{N} x_j\right)^2\right]$$

Et $[E(\bar{X})]^2 = \mu^2 = \left(\frac{1}{N}\sum_{j=1}^{N} x_j\right)^2 = \frac{1}{N^2}\left(\sum_{j=1}^{N} x_j\right)^2$

$\Rightarrow$ (**) $var(\bar{X})$

$$= E(\bar{X}^2) - [E(\bar{X})]^2$$

$$= \frac{1}{nN(N-1)}\left[(N-n)\sum_{j=1}^{N} x_j^2 + (n-1)\left(\sum_{j=1}^{N} x_j\right)^2\right] - \frac{1}{N^2}\left(\sum_{j=1}^{N} x_j\right)^2$$

on distribue et on regroupe ces 2 termes

$$= \frac{N-n}{nN(N-1)}\sum_{j=1}^{N} x_j^2 + \underbrace{\left(\frac{n-1}{nN(N-1)} - \frac{1}{N^2}\right)}_{\frac{n-N}{nN^2(N-1)}}\left(\sum_{j=1}^{N} x_j\right)^2$$

$$= \frac{N-n}{n(N-1)} \left[\underbrace{\frac{1}{N} \sum_{j=1}^{N} x_j^2}_{\text{moyenne des carrés pour la population}} - \underbrace{\left(\frac{1}{N} \sum_{j=1}^{N} x_j \right)^2}_{\text{carré de la moyenne de la population}} \right]$$

$$= \frac{N-n}{n(N-1)} \; \sigma^2$$

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

N.B. si $N \gg n$ $\Rightarrow$ $\frac{N-n}{N-1} \longrightarrow 1$ car n négligeable devant N

$$\Rightarrow \left[Var(\bar{X}) \right]_{EAS} = \left[Var(\bar{X}) \right]_{EASR}$$