Estimation par intervalle de confiance pour la (vraie) moyenne μ

On suppose que l'on dispose d'un échantillon de loi quelconque avec $E[X_i]=\mu$ et $Var(X_i)=\sigma^2$ $\forall i=1,\dots,n$

Par le théorème central limite, on a que $\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to+\infty]{} \mathcal{N}(\mu, \frac{\sigma^2}{n})$

Pour ~~n~~ "assez grand", on peut dire que : (vrai pour n petit si $X_i \sim \mathcal{N}(\mu,\sigma^2)$)

$$P\left(\overline{X} - \underbrace{z_{1-\alpha/2}}_{\uparrow} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

quantile d'ordre $1-\frac{\alpha}{2}$ d'une loi $\mathcal{N}(0;1)$

Le problème est que σ^2 apparait dans les bornes... tout en étant inconnu !!

On remplace les $(X_i)_i$ par des réalisations $(x_i)_i$ et σ^2 par un estimateur sans biais : $\hat{\sigma}^2 = s_{n-1}^2 = \frac{1}{n-1}\sum_i (x_i-\bar{x})^2$

$\hookrightarrow$ $\left[\bar{x} - z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \ ; \ \bar{x} + z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}\right]$ devient l'intervalle de confiance

---

Si $X_i \sim \mathcal{N}(\mu,\sigma^2)$, on utilise la loi de Student.

Soit (i) Z une va de loi $\mathcal{N}(0;1)$ et (ii) U une va indépendante de Z et distribuée selon une loi $\chi^2(k)$. Par définition $T = \frac{Z}{\sqrt{U/k}}$ suit une loi de Student à k degrés de liberté.

On utilise cela pour l'intervalle de confiance de l'estimateur de la moyenne μ d'une v.a. normale de variance σ^2 inconnue :

$$\left[\bar{x} - t^{n-1}_{1-\alpha/2}\sqrt{\frac{S^2_{n-1}}{n}} \; ; \; \bar{x} + t^{n-1}_{1-\frac{\alpha}{2}}\sqrt{\frac{S^2_{n-1}}{n}}\right]$$

où t^k_q est le quantile d'ordre q d'une loi de Student à k degrés de liberté.

En effet on a que : $\displaystyle\sum_i \left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n)$

Voir cours de proba 1ère année

$$\sum_i \left(\frac{X_i-\mu}{\sigma}\right)^2 = \frac{1}{\sigma^2}\sum_i (X_i-\mu)^2$$

$$= \frac{1}{\sigma^2}\sum_i \underbrace{\left[(X_i-\bar{X})+(\bar{X}-\mu)\right]^2}_{=(X_i-\bar{X})^2+2(X_i-\bar{X})(\bar{X}-\mu)+(\bar{X}-\mu)^2}$$

$$= \sum_i \left(\frac{X_i-\bar{X}}{\sigma}\right)^2 + \frac{n}{\sigma^2}(\bar{X}-\mu)^2 \quad \text{car } \sum_i (X_i-\bar{X}) = 0$$

$$= \frac{(n-1)\,S^2_{n-1}}{\sigma^2} + \underbrace{\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2}$$

or $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0;1) \Rightarrow \boxed{\dfrac{(n-1)\,S^2_{n-1}}{\sigma^2} \sim \chi^2(n-1)}$

(au passage, on obtient ici un intervalle de confiance sur σ^2)

Or $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0;1)$ et $\dfrac{(n-1)S^2_{n-1}}{\sigma^2} \sim \chi^2(n-1)$ donc d'où

$$\text{donc } \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \cdot \frac{1}{\sqrt{\frac{(n-1)\,S^2_{n-1}}{\sigma^2}/(n-1)}} \sim \text{Student}(n-1)$$

$= \dfrac{\bar{X}-\mu}{\dots}$ et on a bien l'intervalle annoncé ! $\square$