TENDANCE CENTRALE (C5, F3)
(06 / 10 / 2019, © Monfort, Dicostat2005, 2005-2019)

Posséder une **tendance centrale** est la propriété pour une **variable aléatoire**, un **échantillon**, une **loi de probabilité** ou un **processus** de prendre ses valeurs, ou de se concentrer (cf **concentration**), plus ou moins intensément au **voisinage** d'une valeur (ou d'une région) particulière : cette valeur (ou cette région) est aussi appelée **caractéristique de tendance centrale** (de la variable, de l'échantillon, de la loi ou du processus) (cf **centralité**, **concentration**, **partie centrale**, **valeur centrale**, **variable centrée**) (cf aussi **région de confiance**).

(i) En pratique, la notion concerne souvent des **variables numériques** (scalaires aussi bien que vectorielles) : la tendance centrale est alors classiquement mesurée à l'aide de l'**espérance mathématique** ou **moyenne** théorique (resp de la **moyenne empirique**), de la **médiane** théorique (resp de la **médiane empirique**) ou du **mode** théorique (resp d'un **estimateur** du mode).

Pour une **variable qualitative**, le **mode** peut encore être considéré (ou défini) comme **caractéristique** de tendance centrale ; il est « égal » à la modalité de la variable possédant la plus grande **fréquence** (ou masse).

(ii) Pour une variable numérique, des indicateurs de **dispersion** (**écart-type**, **écart absolu moyen**, etc) permettent de préciser la « **proximité** » moyenne des valeurs (resp des **observations**) autour de la tendance centrale.

(iii) A titre d'exemple, soit $(\Omega, \mathcal{T}, P)$ un **espace probabilisé**, $\xi : \Omega \mapsto \mathbf{R}$ une **va** de **loi** P^{ξ}, $X = (X_1 ,..., X_N)'$ un échantillon indépendant et équidistribué selon P^{ξ} (cf **échantillon iid**) et $\alpha \in \mathbf{R}$ une valeur.

Une **mesure de proximité** de ξ par rapport à α est donnée par le **moment centré** d'ordre 2 (ou sa racine carrée) (cf **écart quadratique moyen**) :

(1) $\mu_{2,\alpha} = E (\xi - \alpha)^2 = \int (x - \alpha)^2 \, dP^{\xi} (x)$.

La décomposition usuelle $E (\xi - \alpha)^2 = E (\xi - E\,\xi + E\,\xi - \alpha)^2 = E (\xi - E\,\xi)^2 + (E\,\xi - \alpha)^2 = V\,\xi + (E\,\xi - \alpha)^2$ montre que ce moment prend une valeur minimum (unique) pour $\alpha^{\sim} = E\,\xi$: cette valeur n'est autre que la variance $V\,\xi = \mu_2 = \sigma^2$. L'écart-type mesure donc la proximité moyenne de ξ autour de son espérance $E\,\xi$.

La version empirique de l'exemple précédent consiste à minimiser :

(2) $\mu_{2,\alpha} (N) = E_N (\xi - a)^2 = \int (x - a)^2 \, dP_N (x)$

pr à a, où P_N désigne la **loi empirique** associée à X. La valeur de a qui minimise $\mu_{2,\alpha} (N)$ est $E_N\,\xi = \overline{X} = e_N' X / e_N' e_N$ (moyenne empirique) et la valeur minimum obtenue n'est autre que $V_N\,\xi = X' P X / e_N' e_N$ (**variance empirique**).

L'exemple précédent s'étend à un **vecteur aléatoire** (va vectorielle) ou à une **distance** donnée par un **moment absolu** d'ordre $j > 2$ (même démarche).