QUEUE D'UNE LOI, D'UNE VARIABLE ALÉATOIRE (C1, C4, C9)
(17 / 01 / 2020, © Monfort, Dicostat2005, 2005-2020)

La **queue** d'une variable aléatoire ou d'une loi de probabilité (**queue de distribution**) correspond, schématiquement, à une région (de l'espace des valeurs) qui est complémentaire d'une **partie centrale** de cette va (ou de cette lp) (cf **centralité**). Elle est notamment associée à l'étude des **valeurs extrêmes**, à une **région critique d'un test**, à l'existence de **moments**, etc.

(i) Soit $(\Omega, \mathcal{T}, P)$ un **espace probabilisé**, $(\mathcal{X}, \mathcal{B})$ un **espace mesurable** et $\xi : \Omega \mapsto \mathcal{X}$ une **va** dont la **loi de probabilité** est P^{ξ} et $C \in \mathcal{B}$ une **partie centrale** associée à P^{ξ} (ou à ξ).

On appelle **queue**, ou **extrémité**, de P^{ξ} (ou de ξ) la partie C^c complémentaire de C.

On peut donc définir une **famille** de queues relatives à une même loi P^{ξ} lorsque la partie centrale C varie.

(ii) On dit que P^{ξ} est une **loi à queue épaisse** lorsque, C étant une partie centrale donnée, on a :

(1) $\quad P^{\xi}(C^c) / P^{\xi}(C) \gg 0.$

Ceci est le cas eg de la **loi de CAUCHY** $\mathcal{C}(a, b)$.

De même, on dit qu'une loi P^{ξ} possède une queue C^c (globalement) plus épaisse que celle d'une loi Q^{ξ} ssi :

(2) $\quad P^{\xi}(C^c) > Q^{\xi}(C^c).$

Ceci implique que les **supports** de ces lois sont inclus dans le même espace de valeurs.

Les lp à queues épaisses jouent un rôle important dans les problèmes de **robustesse** et dans l'étude des **aberrations** (observations atypiques), ainsi que dans les questions de **troncature** des lois ou de **censure** des échantillons. A titre d'exemple, on peut tronquer une loi à partir d'une queue C^c donnée en transformant cette loi en sorte qu'elle ne « charge » que la partie centrale C.

(iii) D'un point de vue terminologique, on dit que C^c est une **queue** de P^{ξ} ou de ξ. Par extension, une partie centrale $C \in \mathcal{B}$ étant donnée, on appelle aussi queue de P^{ξ} ou de ξ l'ensemble $\xi^{-1}(C^c) = [\xi \in C^c]$, ie $\{\omega \in \Omega : \xi(\omega) \in C^c\}$ (**unités statistiques** sur lesquelles on observe ξ).

On dit aussi parfois (par abus de language) que la masse $P^{\xi}(C^c) = 1 - P^{\xi}(C)$ elle-même est la **queue** de la loi P^{ξ}.

(iv) En pratique, une partie centrale C contient souvent (ou est définie par) une **caractéristique** de centralité ou un **paramètre de position** de P^{ξ} (ou de ξ), eg :

(a) $E\,\xi \in C$ si $\xi \in L_{RK}{}^1$ $(\Omega, \mathcal{T}, P)$ (**espérance mathématique**) ;

(b) ou $S\,\xi \in C$ (**mode**) ;

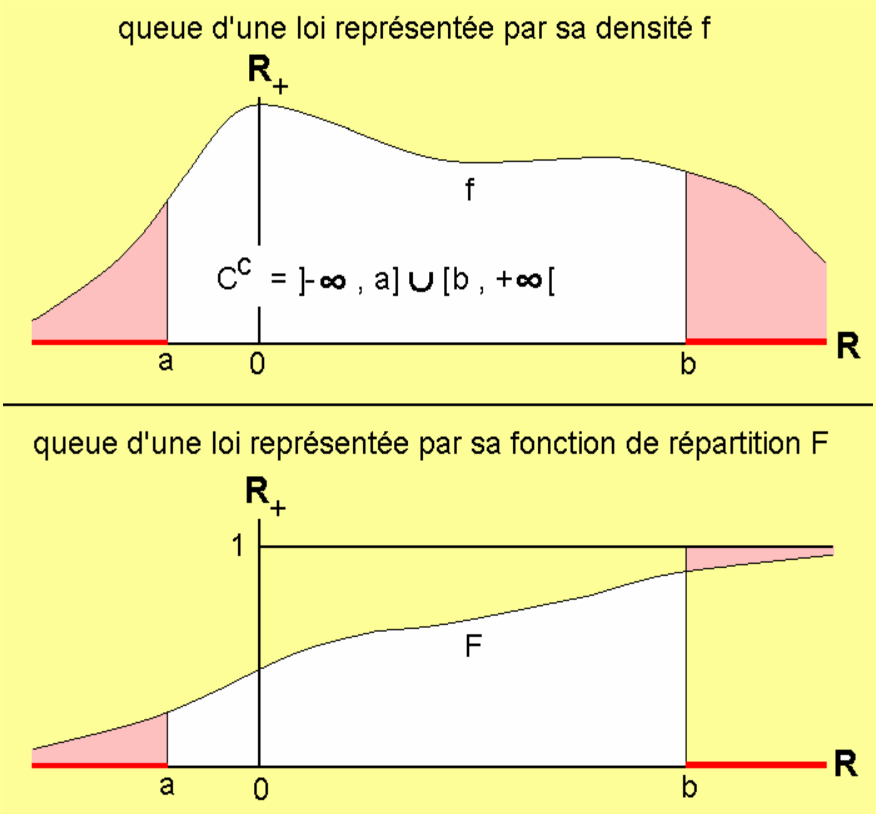
(c) ou encore $Q_{1/2}\,\xi \in C$ (si $\mathcal{X}$ = **R**) (**médiane**).

La queue de la distribution P^{ξ} est généralement « éloignée » de cette caractéristique (ceci dépend aussi de la **concentration** de P^{ξ}) : eg (a) région critique associée à un test d'hypothèses, ou (b) zones de « fortes » fréquences d'observations aberrantes (queues épaisses, **mélanges légaux**).

(v) Dans le cas scalaire, où $\mathcal{X}$ = **R**, une **queue bilatérale** de P^{ξ} est souvent de la forme :

(3) $C^c = C_a \cup C_b =]-\infty, a] \cup [b, +\infty[$, avec $a < 0$ et $b > 0$,

l'ensemble $C_a =]-\infty, a]$ étant appelé **queue à gauche** et l'ensemble $C_b = [b, +\infty[$ **queue à droite** de P^{ξ} (cf schéma ci-après).



On peut ainsi considérer des situations où $|a| >> 0$ et $|b| >> 0$, ou encore étudier le **comportement asymptotique** de la **densité** $f = dP^{\xi} / d\lambda_1$ au voisinage de $-\infty$ ou de $+\infty$, ce qui permet de préciser l'ordre de décroissance des queues.

(vi) On appelle parfois :

(a) **queue** de P^{ξ} le nombre $Q = \int \mathbf{1}(C^c)(x)\, dP^{\xi}(x)$;

(b) **queue à gauche** le nombre $Q_a = \int \mathbf{1}(C_a)(x)\, dP^{\xi}(x)$;

(c) **queue à droite** le nombre $Q_b = \int \mathbf{1}(C_b)(x)\, dP^{\xi}(x)$.

où l'on note $\mathbf{1}(B)$ la **fonction indicatrice** d'une partie B de **R**.

En particulier, on montre eg que :

(4) $\quad Q_b = b \,.\, P([\xi \geq b]) + \int_D P([\xi \geq x])\, dx,\quad$ où $D = \mathbf{R} \setminus C_b$.

Autrement dit :

(5) $\quad Q_b = b \,.\, P^{\xi}(C_b) + \int \mathbf{1}_{[b, +\infty[}(x)\, P^{\xi}(C_x)\, dx.$

(vii) Un problème pratique est celui de l'**approximation** **d'une queue**.

Ainsi, dans le cas gaussien, on cherche à approximer $p = Q(x) = \int d\mathcal{N}(0,1)(x) = (2\pi)^{-1/2} \int x_p\, (1/2) \exp(-x^2)\, dx$, où $x_p = Q^{-1}(p)$.

En supposant eg que $p \leq 1/2$, l'**approximation de H.C. HAMAKER** consiste à remplacer x par $x^{\#} = 1{,}238\, a\, (1 + 0{,}2\, a)$, avec $a = R - \mathrm{Log}\, \{4\, \pi\, (1-\pi)\}$, et p par $p^{\#} = (1/2)\, (1 - (1 - e^{-b(2)}))$, avec $b = 0{,}806$ et $b(2)$ désigne b^2.