Autour de l'inégalité de Bienaymé-Tchebychev

Charles SUQUET

25 novembre 2019

1 Trois démonstrations de l'inégalité

On propose trois démonstrations de l'inégalité de Bienaymé-Tchebychev.

1. La démonstration moderne telle qu'on pourrait l'enseigner en L3. On verra pourquoi elle n'est sans doute pas utilisable en terminale.
2. Une démonstration niveau Terminale qui s'inspire du texte initial de Bienaymé (1853), avec des notations plus lisibles de nos jours.
3. Une preuve graphique presque muette (à part les légendes des figures et la conclusion).

Dans chacune des 3 situations, pour simplifier les notations, on utilise en fait sans le dire l'inégalité de Markov à l'ordre 1 qui ne figure pas au programme mais qui est de toutes façons sous-jacente à l'inégalité de Bienaymé-Tchebychev.

Cela pose d'ailleurs la question de la définition de la variance qui n'est pas donnée (sauf erreur de ma part) dans le programme officiel. La vraie définition est

$$\operatorname{Var} X = \mathbf{E}\big((X - \mathbf{E}X)^2\big). \tag{1}$$

La formule suivante

$$\operatorname{Var} X = \sum_{j=1}^{k}(x_j - \mathbf{E}X)^2 P(X = x_j) = \sum_{x \in X(\Omega)} (x - \mathbf{E}X)^2 P(X = x), \tag{2}$$

où l'on a noté $X(\Omega) = \{x_1, \dots, x_k\}$ l'ensemble fini de toutes les valeurs prises par X, n'est qu'une formule de calcul qu'il conviendrait de justifier à partir de la définition (1).

Pour ce faire on sera amené à poser

$$Y = (X - \mathbf{E}X)^2,$$

et à remarquer que

$$\operatorname{Var} X = \mathbf{E}Y = \sum_{y \in Y(\Omega)} yP(Y = y).$$

Ensuite on note que pour chaque $y \in Y(\Omega)$ il y a un ou deux $x \in X(\Omega)$ tels que $(x - \mathbf{E}X)^2 = y$. S'il y en a un seul, notons le x, alors $P(Y = y) = P(X = x)$ et $yP(Y = y) = (x - \mathbf{E}X)^2 P(X = x)$. S'il y en a deux, notés x' et x'', $P(Y = y) = P(X = x') + P(X = x'')$ et $yP(Y = y) = (x' - \mathbf{E}X)^2 P(X = x') + (x'' - \mathbf{E}X)^2 P(X = x'')$. En recollant tous les morceaux on obtient la formule (2). Ceci étant un peu pénible et chronophage, il me semble préférable d'utiliser Y dans chacune des démonstrations ci-dessous.

Notons aussi pour tout réel $t \geqslant 0$ l'équivalence

$$|X - \mathbf{E}X| \geqslant t \quad \text{si et seulement si} \quad Y \geqslant t^2,$$

d'où

$$P(|X - \mathbf{E}X| \geqslant t) = P(Y \geqslant t^2). \tag{3}$$

Remarque 1 (n'engageant que l'auteur de ces lignes)**.** En relisant le B.O. je vois que la variance de X est notée $V(X)$. Cela me paraît aussi pertinent que de noter $C(\alpha)$ pour $\cos\alpha$. Mais bon le texte officiel étant sacré, je laisse au lecteur le soin de rectifier mes notations.

1.1 La démonstration moderne

On obtient l'inégalité de Bienaymé-Tchebychev grâce aux minorations suivantes valides pour tout $t > 0$ (les justifications sont données après les formules).

$$\begin{align}
\operatorname{Var} X = \mathbf{E}\big((X - \mathbf{E}X)^2\big) &\geqslant \mathbf{E}\big((X - \mathbf{E}X)^2 \mathbf{1}_{\{|X - \mathbf{E}X| \geqslant t\}}\big) \tag{4}\\
&\geqslant \mathbf{E}\big(t^2 \mathbf{1}_{\{|X - \mathbf{E}X| \geqslant t\}}\big) \tag{5}\\
&= t^2 \mathbf{E}\big(\mathbf{1}_{\{|X - \mathbf{E}X| \geqslant t\}}\big) \tag{6}\\
&= t^2 P\big(|X - \mathbf{E}X| \geqslant t\big). \tag{7}\\
& \tag{8}
\end{align}$$

On a ainsi montré que

$$\operatorname{Var} X \geqslant t^2 P\big(|X - \mathbf{E}X| \geqslant t\big).$$

En lisant cette formule de droite à gauche et en divisant par $t^2 > 0$, on obtient l'inégalité de Bienaymé-Tchebychev :

$$\forall t > 0, \quad P\big(|X - \mathbf{E}X| \geqslant t\big) \leqslant \frac{\operatorname{Var} X}{t^2}. \tag{9}$$

Justifications.
(4) et (5) : on utilise la *croissance* de l'espérance ($0 \leqslant Y \leqslant Z \Rightarrow 0 \leqslant \mathbf{E}Y \leqslant \mathbf{E}Z$).
(6) : par linéariré de l'espérance.
(7) : pour tout évènement A, $\mathbf{E}(\mathbf{1}_A) = P(A)$, ce qui revient au calcul de l'espérance d'une v.a. de Bernouilli.

Sauf erreur de ma part, la croissance de l'espérance ne figure pas dans le programme.

1.2 Démonstration basée sur l'idée de Bienaymé

On effectue essentiellement les mêmes minorations que ci-dessus mais en travaillant directement sur les sommes pondérées permettant de calculer la variance. Pour éviter le recours à la notation $\sum$ et aux indices, on peut les présenter sous forme de tableau. Si on choisit cette démonstration, il pourra être utile de commencer par traiter « à la main » un exemple où Y ne prend que 3 ou 4 valeurs.

		Somme des termes de la forme	pour y vérifiant
$\operatorname{Var} X$	$=$	$yP(Y = y)$	$y \in Y(\Omega)$
	$\geqslant$	$yP(Y = y)$	$y \geqslant t^2$
	$\geqslant$	$t^2P(Y = y)$	$y \geqslant t^2$
	$=$	$t^2P(Y \geqslant t^2) = t^2P(\|X - \mathbf{E}X\| \geqslant t)$	

et on termine comme dans la première démonstration.

1.3 Preuve (presque) muette

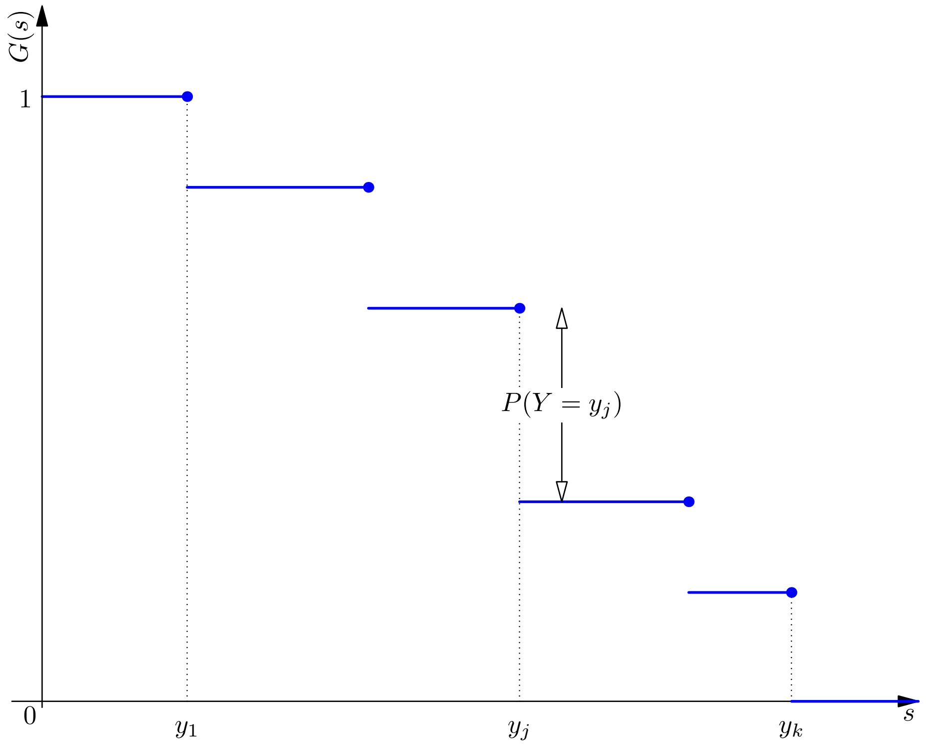

FIGURE 1 – Fonction $G : s \mapsto P(Y \geqslant s)$, Y est la v.a. positive $(X - \mathbf{E}X)^2$

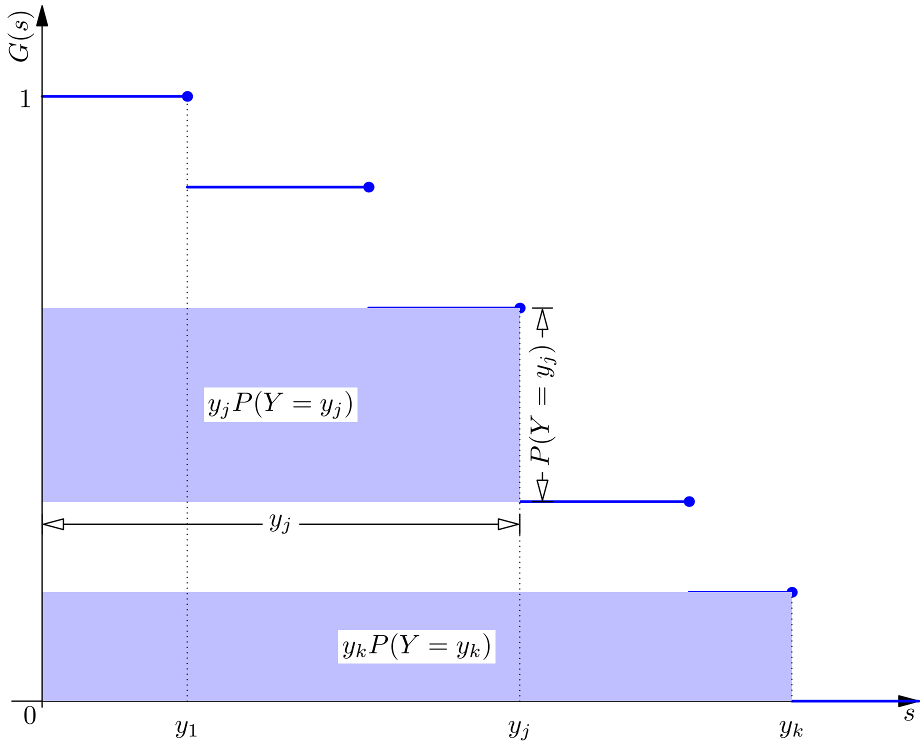

FIGURE 2 – Découpage en tranches de l'aire sous l'escalier G

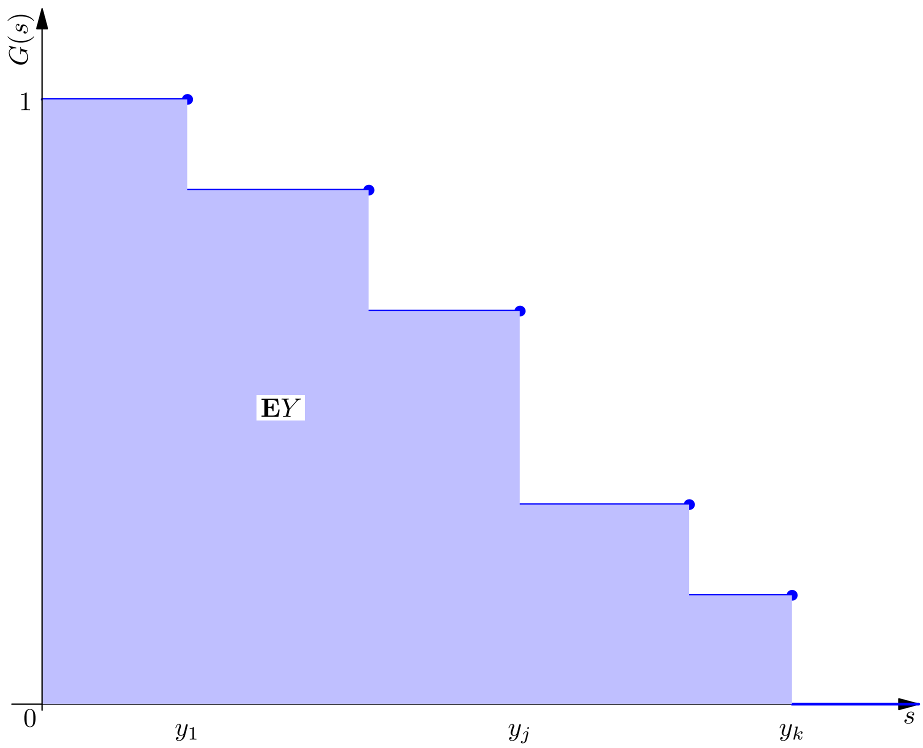

FIGURE 3 – L'espérance de Y est l'aire sous l'escalier G

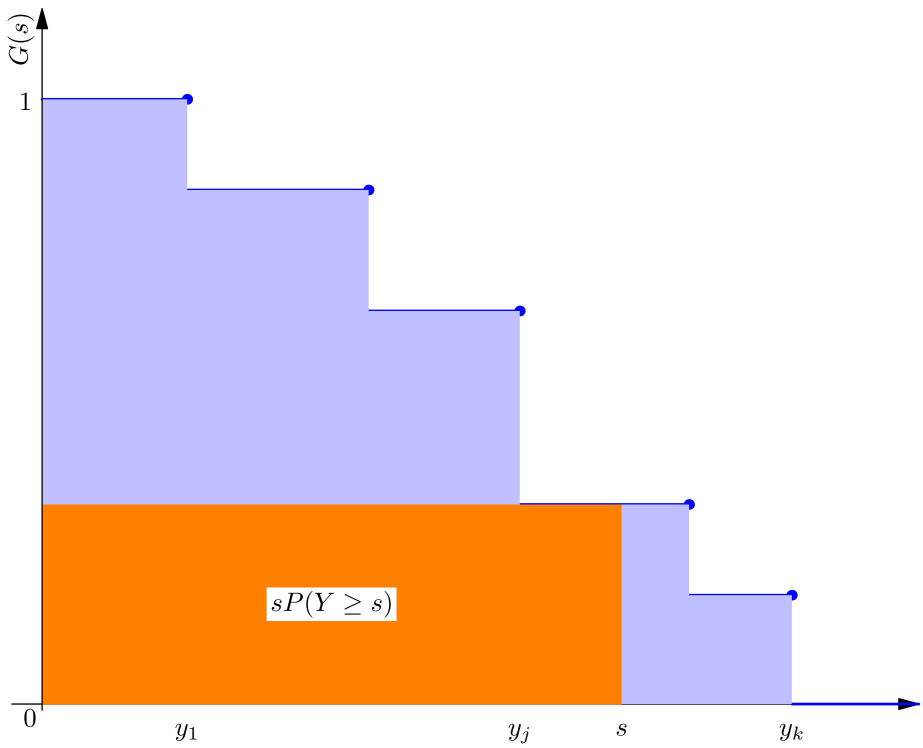

FIGURE 4 – $sP(Y \geqslant s) \leqslant \mathbf{E}Y$, d'où avec $s = t^2$, $P(|X - \mathbf{E}X| \geqslant t) \leqslant \frac{\operatorname{Var} X}{t^2}$