VARD : preuve du théorème 49 (linéarité de l'espérance)

On démontre le théorème uniquement dans le cas où X et Y sont deux VARD finies. Dans ce cas, la variable aléatoire $\lambda X + \mu Y$ prend un nombre fini de valeurs (les $\lambda x + \mu y$ avec $x \in X(\Omega)$ et $y \in Y(\Omega)$), donc elle admet une espérance donnée par

$$\begin{aligned}
\mathbb{E}(\lambda X + \mu Y) &= \sum_{z \in (\lambda X + \mu Y)(\Omega)} z\, \mathbb{P}(\lambda X + \mu Y = z) \\
&= \sum_{z \in (\lambda X + \mu Y)(\Omega)} z \left[\sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ \lambda x + \mu y = z}} \mathbb{P}((X = x) \cap (Y = y)) \right] \\
&= \sum_{z \in (\lambda X + \mu Y)(\Omega)} \left[\sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ \lambda x + \mu y = z}} (\lambda x + \mu y) \mathbb{P}((X = x) \cap (Y = y)) \right] \\
&= \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} (\lambda x + \mu y)\, \mathbb{P}((X = x) \cap (Y = y)) \\
&= \lambda \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} x\, \mathbb{P}((X = x) \cap (Y = y)) + \mu \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} y\, \mathbb{P}((X = x) \cap (Y = y)) \\
&= \lambda \sum_{x \in X(\Omega)} x \sum_{y \in Y(\Omega)} \mathbb{P}((X = x) \cap (Y = y)) + \mu \sum_{y \in Y(\Omega)} y \sum_{x \in X(\Omega)} \mathbb{P}((X = x) \cap (Y = y)) \\
&= \lambda \sum_{x \in X(\Omega)} x\, \mathbb{P}(X = x) + \mu \sum_{y \in Y(\Omega)} y\, \mathbb{P}(Y = y) \\
&= \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y),
\end{aligned}$$

ce qu'il fallait démontrer.

Dans ce qui précède, la deuxième et la troisième égalité sont valides parce que tout $z \in (X + \lambda Y)(\Omega)$ est de la forme $\lambda x + \mu y$ avec $x \in X(\Omega)$ et $y \in Y(\Omega)$. Les réarrangements de sommes (linéarité, interversions et sommation par paquets) sont permis par le fait que toutes les sommes mobilisées sont finies. Enfin, on a utilisé la σ-additivité de $\mathbb{P}$ (c'est-à-dire le fait que la probabilité d'une union disjointe est la somme des probabilités) dans la deuxième et la sixième égalité.

La preuve est la même dans le cas où X et Y prennent un nombre infini de valeurs, mais il faut ajouter un certain nombre de considérations de convergence permettant les diverses manipulations effectuées, ce dont nous nous dispensons ici. $\square$