Énoncé

Soit X une variable aléatoire réelle discrète. On suppose que $X \geq 0$ et $\mathbb{E}(X) = 0$.
Montrer que : X = 0 presque sûrement i.e $\mathbb{P}(\{X = 0\}) = 1$.

Correction

Pourquoi le résultat semble vrai ?

Tout d'abord, comme X est positive, la moyenne de ses valeurs (si on répète l'expérience aléatoire un grand nombre de fois) $\mathbb{E}(X)$ est positive.
Or, on a par hypothèse que cette moyenne est nulle.
Intuitivement, on peut se dire que X ne peut pas prendre trop souvent des valeurs strictement positives car sinon $\mathbb{E}(X)$ serait strictement positive étant donné qu'il n'y a pas de termes négatifs pour compenser la moyenne. Donc il est raisonnable de penser que X est presque tout le temps nulle : $\mathbb{P}(X = 0) = 1$.

Première preuve

Etape 1 :
Une piste naturelle à explorer serait de partir de l'égalité :

$$\mathbb{E}(X) = 0$$

On introduit alors des notations pour pouvoir réécrire cette égalité :
L'ensemble des valeurs prises par X étant supposé au plus dénombrable, on peut le noter ainsi : $\{x_n, n \in \mathbb{N}\} \subset \mathbb{R}_+$. Rappelons qu'un ensemble est au plus dénombrable s'il est **fini** ou **dénombrable**. Si $X(\Omega)$ est dénombrable (c'est-à-dire en bijection avec $\mathbb{N}$) , la notation est claire. Si $X(\Omega)$ est fini, on a un nombre fini de $x_n, n \in \mathbb{N}$ **distincts** et $\{x_n, n \in \mathbb{N}\}$ se réécrit tout simplement $\{x_{n_1}, x_{n_2}, \ldots, x_{n_p}\}$ avec $p \in \mathbb{N}^*$ le cardinal de $X(\Omega)$ et $(n_1, n_2, \ldots, n_p) \in \mathbb{N}^p$ des indices séléctionnés donnant les valeurs distinctes de la suite $(x_n)_{n\in\mathbb{N}}$. Cette écriture correspond à la notation usuelle d'un ensemble fini non vide.

L'égalité $\mathbb{E}(X) = 0$ se réécrit par définition de l'espérance :

$$\sum_{n\in\mathbb{N}} \{x_n \mathbb{P}(X = x_n)\} = 0$$

Comme tous les termes de la famille $(x_n \mathbb{P}(X = x_n))_{n\in\mathbb{N}}$ sont positifs, on en déduit :

$$\forall n \in \mathbb{N} \quad x_n \mathbb{P}(X = x_n) = 0$$

$$\Leftrightarrow \boxed{\forall n \in \mathbb{N} \quad x_n = 0 \quad \text{ou} \quad \mathbb{P}(X = x_n) = 0} \quad *$$

Interprétation du résultat

Cette assertion signifie que pour $n \in \mathbb{N}$ si x_n est non nul, la probabilité que l'événement $\{X = x_n\}$ se réalise est nulle.
Donc en dehors de l'événement $\{X = 0\}$, tous les autres événements de la forme $\{X = x_n\}$, $n \in \mathbb{N}$ ont aucune chance de se produire. Pourtant, après chaque réalisation de l'expérience, X prend une valeur, l'évènement $X \in \{x_n, n \in \mathbb{N}\}$ est réalisé. La seule possibilité est que presque sûrement, à chaque répétition de l'expérience aléatoire, c'est l'évènement $\{X = 0\}$ qui est réalisé, impliquant la réalisation de $X \in \{x_n, n \in \mathbb{N}\}$.

Etape 2 :

Pour conclure formellement, on va utiliser cet axiome d'une probabilité :

$$\mathbb{P}(\Omega) = 1$$

qui peut se réécrire :

$$\mathbb{P}(X \in \{x_n, n \in \mathbb{N}\}) = 1$$

(On rappelle que par définition, $\quad \mathbb{P}(X \in \{x_n, n \in \mathbb{N}\}) := \mathbb{P}(X^{-1}(\{x_n, n \in \mathbb{N}\}))$)

$$\Leftrightarrow \mathbb{P}(\bigcup_{n \in \mathbb{N}} \{X = x_n\}) = 1$$

Les évènements $\{X = x_n\}$ et $\{X = x_p\}$ étant incompatibles dès que n $n \neq p$, par σ-additivité :

$$\Leftrightarrow \sum_{n \in \mathbb{N}} \mathbb{P}\{X = x_n\} = 1$$

Explication pour la suite

Là on a presque fini, il suffit d'éliminer les $x_n, n \in \mathbb{N}$ non nuls de la somme pour récupérer $\mathbb{P}(X = 0) = 1$.

Etape 3 :

Notons $A := \{n \in \mathbb{N} | x_n > 0\}$.
Alors $A^c = \{n \in \mathbb{N} | x_n = 0\}$ et $\{X = 0\} = \bigcup_{n \in A^c} \{X = x_n\}$.

$$\sum_{n \in A \sqcup A^c} \mathbb{P}\{X = x_n\} = 1$$

Or d'après la propriété $*$,

$$\forall n \in A \quad \mathbb{P}(X = x_n) = 0$$

Donc

$$\sum_{n \in A^c} \mathbb{P}\{X = x_n\} = 1$$

On réutilise la σ-additivité dans l'autre sens :

$$\mathbb{P}(\bigcup_{n\in A^c}\{X = x_n\}) = 1$$

$$\boxed{\mathbb{P}(\{X = 0\}) = 1}$$

Remarque

La première preuve est compréhensible mais n'est pas facile à rédiger rigoureusement. Pour simplifier les écritures, j'aurais voulu supposer que $x_0 = 0$ mais celà demande quelques justifications :

1) Démontrer que 0 est atteint par X (par l'absurde avec $\mathbb{E}(X) = 0$)

2) Justifier que lorsqu'on note $\{x_n, n \in \mathbb{N}\}$ les valeurs prises par X, on peut supposer que $x_0 = 0$. Il faut pour cela distinguer les cas $X(\Omega)$ fini ou dénombrable. Puis justifier dans chacun des cas qu'on peut ranger dans l'ordre strictement croissant les $x_n, n \in \mathbb{N}$ et donc quitte à renuméroter les $x_n, n \in \mathbb{N}$, supposer que $x_0 = 0$.

Deuxième preuve

Proposons une autre approche similaire.

Etape 1 :

On aurait pu penser à travailler avec l'évènement contraire $\{X = 0\}^c$. En effet, connaître la probabilité de l'évènement $\{X = 0\}$ est équivalent à connaître celle de $\{X = 0\}^c$, d'après la relation fondamentale :

$$\mathbb{P}(\{X = 0\}) + \mathbb{P}(\{X = 0\}^c) = 1$$

On a :

$$\{X = 0\}^c = \{X \neq 0\} = \{X > 0\} \quad \text{car } X \text{ ne prend que des valeurs positives.}$$

Donc montrer que $\mathbb{P}(\{X = 0\}) = 1$ est équivalent à montrer que $\mathbb{P}(\{X > 0\}) = 0$.

Etape 2 :

En reprenant les notations de la première résolution, on écrit :

$$\{X > 0\} = \bigcup_{n\in A}\{X = x_n\}$$

Avec d'après la propriété *,

$$\forall n \in A \quad \mathbb{P}(X = x_n) = 0$$

On a :

$$\boxed{\mathbb{P}(X > 0) = \mathbb{P}(\bigcup_{n \in A} \{X = x_n\})}$$

Etape 3 :
Comme A est au plus dénombrable, d'après **l'inégalité de Boole** : (on a même encore l'égalité ici car les évènements sont disjoints et on peut utiliser la σ-additivité)

$$\mathbb{P}(X > 0) \leq \sum_{n \in A} \mathbb{P}\{X = x_n\}$$

$$\mathbb{P}(X > 0) \leq \sum_{n \in A} 0$$

$$\mathbb{P}(X > 0) \leq 0$$

Donc

$$\boxed{\mathbb{P}(X > 0) = 0}$$

Remarque
Cette preuve est plus efficace et plus facile à rédiger que la première. Une fois l'assertion : "$\forall n \in \mathbb{N} \quad x_n = 0 \quad \text{ou} \quad \mathbb{P}(X = x_n) = 0$" obtenue, on peut naturellement décider de travailler sur l'évènement contraire, plus facile à manipuler et calculer sa probabilité.

Troisième preuve

Enfin, voyons une approche totalement différente en raisonnant encore sur l'évènement contraire.

Etape 1 :
Comme une probabilité est positive, $\mathbb{P}(\{X > 0\}) = 0$ est équivalent à :

$$\mathbb{P}(\{X > 0\}) \leq 0$$

En sachant que par hypothèse $\mathbb{E}(X) = 0$,
L'inégalité à démontrer se réécrit :

$$\mathbb{P}(\{X > 0\}) \leq \mathbb{E}(X)$$

Cela ressemble beaucoup à **l'inégalité de Markov** :

$$\forall a \in \mathbb{R}_+^* \quad \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

Etant donné l'inclusion $\{X > a\} \subset \{X \geq a\}$, on a :

$$\forall a \in \mathbb{R}_+^* \quad \mathbb{P}(X > a) \leq \mathbb{P}(X \geq a)$$

D'où par transitivé de $(\leq)$, on a l'inégalité :

$$\boxed{\forall a \in \mathbb{R}_+^* \quad \mathbb{P}(X > a) \leq \frac{\mathbb{E}(X)}{a}}$$

Analyse de la difficulté et solution

Malheureusement, on ne peut pas appliquer la dernière inégalité ici ($a = 0$ pour nous). Il faut ruser un peu. L'idée est de réécrire l'évènement $\{X > 0\}$ comme une réunion au plus dénombrable d'évènements de la forme $\{X > a_n\}$ où $n \in \mathbb{N}$ et $a_n > 0$ et d'appliquer **l'inégalité de Boole** pour isoler les $\mathbb{P}(X > a_n)$ puis appliquer **l'inégalité de Markov** pour majorer par 0 la probabilité de chacun de ces évènements.

Etape 2 :

On écrit :

$$\{X > 0\} = \bigcup_{n \in \mathbb{N}} \{X > \frac{1}{2^n}\}$$

En effet, si $\omega \in \Omega$ est tel que $X(\omega) > 0$ ($\omega \in \{X > 0\}$), alors comme $\frac{1}{2^n} \to 0$ quand $n \to +\infty$, on dispose d'un rang $N \in \mathbb{N}$ tel que :

$$\forall n \in \mathbb{N} \quad n \geq N \Rightarrow \frac{1}{2^n} < X(\omega)$$

Donc en particulier $X(\omega) > \frac{1}{2^N}$ et $\omega \in \{X > \frac{1}{2^N}\}$.

Ce qui montre

$$\omega \in \bigcup_{n \in \mathbb{N}} \{X > \frac{1}{2^n}\}$$

Donc

$$\{X > 0\} \subset \bigcup_{n \in \mathbb{N}} \{X > \frac{1}{2^n}\}$$

Réciproquement, soit $n \in \mathbb{N}$.

Si $\omega \in \Omega$ est tel que $X(\omega) > \frac{1}{2^n}$ alors comme $1/2^n > 0$ on a par transitivé $X(\omega) > 0$.

Ce qui montre

$$\{X > 0\} \ni \omega$$

Donc

$$\{X > 0\} \supset \bigcup_{n \in \mathbb{N}} \{X > \frac{1}{2^n}\}$$

Par double inclusion,

$$\{X > 0\} = \bigcup_{n\in\mathbb{N}} \{X > \frac{1}{2^n}\}$$

Remarques :
On aurait pu prendre ici n'importe quelle suite de d'éléments de $\mathbb{R}_+^*$ qui converge vers 0. Les deux seuls arguments qu'on a utilisé concernant $(a_n)_{n\in\mathbb{N}}$ sont la stricte positivité et la convergence vers 0.
Le fait que l'union soit indexée par un ensemble au plus dénombrable est primordial. On pourrait très bien écrire :

$$\{X > 0\} = \bigcup_{x\in\mathbb{R}} \{X > \frac{1}{1+x^2}\}$$

Mais cette indexation par un ensemble infini indénombrable ($\mathbb{R}$) ne permet plus d'appliquer l'inégalité de Boole. De plus, on ne sait pas à priori si l'union indénombrable d'évènements est encore un évènement (i.e un élément de la tribu). L'application $\mathbb{P}$ peut ne pas être définie en cette union !

Etape 3 :
L'inégalité de Boole donne :

$$\mathbb{P}(\bigcup_{n\in\mathbb{N}} \{X > \frac{1}{2^n}\}) \leq \sum_{n\in\mathbb{N}} \mathbb{P}\{X > \frac{1}{2^n}\}$$

On applique enfin l'inégalité de Markov,

$$\forall n \in \mathbb{N} \quad \mathbb{P}\{X > \frac{1}{2^n}\} \leq 2^n \mathbb{E}(X) = 0$$

Donc

$$\mathbb{P}(\bigcup_{n\in\mathbb{N}} \{X > \frac{1}{2^n}\}) \leq \sum_{n\in\mathbb{N}} 0$$

$$\Leftrightarrow \mathbb{P}(\bigcup_{n\in\mathbb{N}} \{X > \frac{1}{2^n}\}) \leq 0$$

$$\Leftrightarrow \mathbb{P}(\bigcup_{n\in\mathbb{N}} \{X > \frac{1}{2^n}\}) = 0$$

$$\Leftrightarrow \boxed{\mathbb{P}(\{X > 0\}) = 0}$$

Remarque
La troisième preuve est la plus efficace mais elle n'est pas très réaliste à moins d'avoir déjà vu une écriture du type $\{X > 0\} = \bigcup_{n\in\mathbb{N}} \{X > \frac{1}{2^n}\}$ et d'en plus penser à l'utiliser ici.

résumé des preuves

Preuve 1 :

1) Récupérer l'info donnée par $\mathbb{E}(X) = 0$: $\forall n \in \mathbb{N} \quad x_n = 0 \quad \text{ou} \quad \mathbb{P}(X = x_n) = 0$
2) Partir de $\mathbb{P}(X \in \{x_n, n \in \mathbb{N}\}) = 1$ et utiliser la σ-additivé :
   $\mathbb{P}(X \in \{x_n, n \in \mathbb{N}\}) = \sum_{n\in\mathbb{N}} \mathbb{P}\{X = x_n\}$
3) Réécrire $\sum_{n\in\mathbb{N}} \mathbb{P}\{X = x_n\} = 1$ en séparant $\mathbb{P}(\{X = 0\})$ du reste des $\mathbb{P}(\{X = x_n\}),\ n \in \mathbb{N}$.

Preuve 2 :

1) Récupérer l'info donnée par $\mathbb{E}(X) = 0$: $\forall n \in \mathbb{N} \quad x_n = 0 \quad \text{ou} \quad \mathbb{P}(X = x_n) = 0$
2) Considérer l'événement complémentaire : $\{X > 0\}$
3) Le réécrire comme une union **dénombrable** d'évènements de probabilité nulle
   $\{X > 0\} = \bigcup_{n\in A}\{X = x_n\}$
4) Appliquer la σ-additivité (licite) ou l'inégalité de Boole à $\bigcup_{n\in A}\{X = x_n\}$

Preuve 3 :

1) Considérer l'événement complémentaire : $\{X > 0\}$
2) Le réécrire comme une union **dénombrable** :$\{X > 0\} = \bigcup\limits_{n\in\mathbb{N}}\{X > \frac{1}{2^n}\}$
3) Appliquer l'inégalité de Boole : $\mathbb{P}(\bigcup\limits_{n\in\mathbb{N}}\{X > \frac{1}{2^n}\}) \leq \sum\limits_{n\in\mathbb{N}} \mathbb{P}\{X > \frac{1}{2^n}\}$
4) Appliquer l'inégalité de Markov : $\forall n \in \mathbb{N} \quad \mathbb{P}\{X > \frac{1}{2^n}\} \leq 2^n\mathbb{E}(X) = 0$
5) Conclure que $\mathbb{P}\{X > 0\} \leq 0$