Variables aléatoires réelles discrètes

On réalise une expérience aléatoire et on considère l'univers $\Omega = \{\text{issues possibles à l'expérience}\}$.
On munit Ω d'une tribu $\mathcal{A}$ de parties de Ω, puis on définit une probabilité P sur cet espace.

1 Loi d'une variable discrète

1.1 Généralités

Rappel : une variable aléatoire discrète sur un espace probabilisable $(\Omega, \mathcal{A})$ est une application de Ω dans $\mathbb{R}$ telle que :

1. pour tout réel x, $(X \leq x) = \{\omega \in \Omega | X(\omega) \leq x\} \in \mathcal{A}$
2. $X(\Omega)$, l'ensembles des valeurs prises par X, est un ensemble fini ou dénombrable.

Exemple 1 :
On effectue une infinité de lancers d'un dé. On introduit la variable aléatoire X qui donne le nombre de lancers nécessaires à l'obtention du premier 1, si l'on obtient 1, et 0 si l'on n'obtient jamais 1.
Alors, introduite ainsi, on a : $X(\Omega) = [\![1, +\infty[\![\cup \{0\} = \mathbb{N}$.
En fait, beaucoup d'énoncés définiront X sans prendre la peine de définir la valeur 0 c'est-à-dire :
soit X la variable aléatoire égale au nombre de lancers nécessaires à l'obtention du premier 1.
En effet : ces deux définitions de X reviennent au même car on a vu dans le chapitre précédent que presque sûrement, on effectuera un nombre fini de lancers puisque d'après le théorème de limite monotone, $P(X = 0) = P(\text{"ne jamais obtenir 1" }) = \lim\limits_{n \to +\infty} P(\text{"ne pas obtenir de 1 en } n \text{ lancers"})=0$.
Donc même avec la première définition de X, X ne prend pas la valeur 0 presque sûrement, et donc on peut considérer que $X(\Omega) = \mathbb{N}^*$ (ce que l'énoncé vous demandera de vérifier).

Remarque Autant il n'est pas toujours possible de définir Ω, autant il est indispensable de préciser $X(\Omega)$ dès qu'une variable aléatoire X est introduite.

`Théorème`
Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A})$ et $X(\Omega)$ l'ensemble des valeurs prises par X. Alors la famille $([X = k])_{k \in X(\Omega)}$ est un s.c.e. appelé système complet d'événements associé à X.

Définition **
Soit X une var discrète sur $(\Omega, \mathcal{A})$. On appelle tribu engendrée par X, la tribu engendrée par les événements $([X = k])_{k \in X(\Omega)}$. Cette tribu contient tous les événements qui sont des intersections finies ou dénombrables ou réunion finies ou dénombrables de ces événements ou de leur complémentaires.

1.2 Loi de probabilité

Définition
Soit X une variable discrète sur $(\Omega, \mathcal{A}, P)$
On appelle loi de probabilité de X (ou loi de X ou distribution de X) l'ensemble des couples
$\{k, P(X = k)), \ k \in X(\Omega)\}$.

Retour exemple 1 : On effectue une infinité de lancers.
X est le nombre de lancers de dés nécessaires à l'obtention du premier 1. $X(\Omega) = \mathbb{N}^*$. Loi de X ?
On introduit pour tout $k \in \mathbb{N}^*$ les événements U_k "obtenir un 1 au k^{ie} lancer".
Alors $(X = 1) = U_1$ et $P(X = 1) = 1/6$. Puis pour tout $k \geq 2$, $(X = k) = \overline{U_1} \cap \overline{U_2} \cap ... \cap \overline{U_{k-1}} \cap U_k$ et par indépendance des lancers, on obtient $P(X = k) = (5/6)^{k-1} * 1/6$.
On vérifie alors que $\sum\limits_{k=1}^{+\infty} P(X = k) = \sum\limits_{k=1}^{+\infty} \frac{1}{6}(\frac{5}{6})^{k-1} = \frac{1}{6} \sum\limits_{k=0}^{+\infty} (\frac{5}{6})^k = \frac{1}{6} \frac{1}{1-5/6} = 1$.

`Remarque 1` : Cela permettrait de montrer d'une autre façon que $P(X = 0) = 0$ (alternative au théorème de limite monotone) puisque $P(X = 0) = 1 - \sum\limits_{k=1}^{+\infty} P(X = k) = 1 - 1 = 0$.

`Remarque 2` : la convergence de la série vient de la propriété de $\sigma-$additivité de P (cf chapitre précédent) : en effet, la famille $(X=k)_{k\in X(\Omega)}$ est formée d'événements deux-à-deux incompatibles, donc la série $\sum\limits_{k\geq 1} P(X=k)$ converge.

`Théorème`
Un ensemble $\{(x_k,p_k),\ k\in I\}$ définit une loi de probabilité d'une variable aléatoire discrète ssi

1. $\forall k \in I\ p_k \geqslant 0$
2. $\sum\limits_{k\in I} p_k = 1$ (sous-entendu dans le cas I infini : la série converge et la somme de la série vaut 1)

Exemple :
$\{(2k+1, \frac{2}{3^{k+1}}), k\in\mathbb{N}\}$ définit la loi de probabilité d'une variable discrète X car $\forall k\in\mathbb{N}$, $\frac{2}{3^{k+1}} \geq 0$ et $\frac{2}{3^{k+1}} = \frac{2}{3}\frac{1}{3^k}$ donc la série $\sum\limits_{k\geq 0}\frac{2}{3^{k+1}}$ converge et sa somme vaut $\frac{2}{3}\sum\limits_{k=0}^{+\infty}\frac{1}{3^k} = \frac{2}{3}\frac{1}{1-1/3} = 1$. Une telle variable prendrait comme valeurs tous les entiers impairs.

1.3 Fonction d'une variable réelle discrète

Comme dans le chapitre précédent, on peut vouloir étudier des variables aléatoires du type X^2, ou ...
`Proposition`
Soit X une variable discrète sur $(\Omega,\mathcal{A},P)$, f une fonction réelle définie sur $X(\Omega)$ et $Y=f(X)$ (autrement dit, $\forall\omega\in\Omega$, $Y(\omega)=f(X(\omega))$).
Alors Y est une variable aléatoire discrète sur $(\Omega,\mathcal{A},P)$ et

1. $Y(\Omega)=\{f(k), k\in X(\Omega)\}$
2. pour tout $y\in Y(\Omega)$, $P(Y=y) = \sum\limits_{k\in X(\Omega)/f(k)=y} P(X=k)$

2 Fonction de répartition

En plus des propriétés vues dans le précédent chapitre, on peut généraliser le cas X variable finie :
`Théorème`
Soit X une variable aléatoire discrète infinie définie sur un espace probabilisé $(\Omega,\mathcal{A},P)$ et
$X(\Omega)=\{x_0,x_1,...x_n,...\}$ avec $x_0<x_1<...<x_n<...$ ($(x_n)_{n\in\mathbb{N}}$ suite strictement croissante).
Alors la fonction de répartition de X, F_X, est une fonction constante par morceaux, car constante sur chaque intervalle $[x_k,x_{k+1}[$. Autrement dit, la fonction de répartition peut être visualisée comme une fonction en escalier "généralisée" (avec une infinité de palliers).

Remarque
Pourquoi l'hypothèse $(x_n)_{n\in\mathbb{N}}$ croissante ? : considérer une variable X de loi
$X(\Omega)=\{\frac{1}{n}, n\in\mathbb{N}^*\}$ et $\forall n\geq 1$, $P(X=\frac{1}{n})=\frac{1}{2^n}$.
La fonction de répartition reste bien sûr constante sur chaque intervalle de type $[\frac{1}{n+1},\frac{1}{n}[$, mais les palliers se "tassent " en 0.

`Théorème admis`
La fonction de répartition caractérise la loi.

En pratique :

1. La connaissance de la loi de X permet la détermination de F_X.
   En effet, dans le cas où $X(\Omega)\subset\mathbb{N}$, pour tout $n\in\mathbb{N}$, $(X\leq n)=\bigcup\limits_{k=0}^{n}(X=k)$ d'où par $\sigma-$additivité, $P(X\leq n)=\sum\limits_{k=0}^{n}P(X=k)$.
2. Réciproquement : la connaissance de la fonction de répartition de X permet la détermination de la loi de X.

En effet dans le cas où $X(\Omega) \subset \mathbb{N}$, pour tout $n \in \mathbb{N}$, $(X \leq n) = (X = n) \cup (X \leq n-1)$ réunion de 2 événements incompatibles, d'où $P(X = n) = P(X \leq n) - P(X \leq n-1) = F_X(n) - F_X(n-1)$

3 Moments d'une variable aléatoire discrète

3.1 Espérance

DÉFINITION
Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, P)$.
On dit que X admet une espérance si $X(\Omega)$ est fini ou si la série $\sum\limits_{k \in X(\Omega)} kP(X = k)$ converge absolument.
On appelle alors espérance de X le réel $E(X) = \sum\limits_{k \in X(\Omega)} kP(X = k)$ (qui est la somme de la série, lorsque $X(\Omega)$ est infini).

Exemple 1 : X est le nombre de lancers de dés nécessaires pour obtenir le chiffre 1.
On a vu que $X(\Omega) = \mathbb{N}^*$ et que $\forall k \in \mathbb{N}^*$, $P(X = k) = (5/6)^{k-1} 1/6$.
Or la série $\sum n x^{n-1}$ (série géométrique dérivée) converge (absolument) pour tout $x \in]-1, 1[$, donc la série $\sum\limits_{n \geq 1} nP(X = n) = \sum\limits_{n \geq 1} \frac{1}{6} n \frac{5}{6}^{n-1}$ converge absolument et X admet une espérance : $E(X) = \frac{1}{6} \frac{1}{(1-\frac{5}{6})^2} = 6$

`Proposition [Formule de transfert]admis`
Soit X une variable discrète sur $(\Omega, \mathcal{A}, P)$, et f une fonction réelle.
La variable aléatoire $f(X)$ admet une espérance si et seulement si $X(\Omega)$ est fini ou si la série $\sum\limits_{k \in X(\Omega)} f(k)P(X = k)$ converge absolument, et alors $E(f(X)) = \sum\limits_{k \in X(\Omega)} f(k)P(X = k)$ (qui est la somme de la série, lorsque $X(\Omega)$ est infini).

Cas particulier : $f(x) = x^2$.

X^2 admet une espérance si $X(\Omega)$ est fini ou si la série $\sum\limits_{k \in X(\Omega)} k^2 P(X = k)$ converge.

(Dans ce cas particulier, inutile de justifier la convergence absolue, car tous les termes sont positifs !)
Et $E(X^2) = \sum\limits_{k \in X(\Omega)} k^2 P(X = k)$.

`Proposition [linéarité de l'espérance]`
Soit X une variable discrète sur $(\Omega, \mathcal{A}, P)$ admettant une espérance et a, b deux nombres réels.
Alors la variable $aX + b$ admet une espérance et $E(aX + b) = aE(X) + b$.

Cette propriété sera complétée l'an prochain par :
Soit X et Y deux variables discrètes sur $(\Omega, \mathcal{A}, P)$ admettant une espérance.
Alors la variable $X + Y$ admet une espérance et $E(X + Y) = E(X) + E(Y)$.

DÉFINITION

1. On dit que la variable X est centrée si $E(X) = 0$.
2. La variable $X - E(X)$ est appelée variable centrée associée à X.

En effet, posons $Y = X - E(X)$ et montrons que $E(Y) = 0$.
Par linéarité de l'espérance, avec $a = 1$ et $b = -E(X) \in \mathbb{R}$, $E(Y) = E(X - E(X)) = E(X) - E(X) = 0$.

3.2 Variance

DÉFINITION
Soit X une var discrète sur $(\Omega, \mathcal{A}, P)$. Sous-réserve d'existence, on appelle

1. moment d'ordre 2 de X le nombre $m_2(X) = E(X^2)$

2. variance de X le nombre $V(X) = E[(X - E(X))^2]$ (mesure l'écart entre X et sa moyenne)
3. écart-type de X le nombre $\sigma(X) = \sqrt{V(X)}$

Remarque

X peut admettre une espérance sans admettre de moment d'ordre 2.
Mais si X admet un moment d'ordre 2, alors X admet une espérance, et donc par contraposée, si X n'admet pas d'espérance alors X n'admet pas non plus de moment d'ordre 2.
Pour le comprendre, soit X une var discrète avec $X(\Omega) = \mathbb{N}$.
Alors l'espérance existe si la série $\sum\limits_{k\geq 0} kP(X = k)$ converge et $E(X^2)$ existe si la série $\sum\limits_{k\geq 0} k^2P(X = k)$. Or, pour tout $k \geq 0$, $k^2 \geq k$ donc (critère de comparaison), si la série de l'espérance diverge, l'autre aussi, et si la série associée au moment d'ordre 2 converge, celle de l'espérance aussi !

`Proposition`

X admet un moment d'ordre 2 ssi X admet une variance et dans ce cas, $V(X) = E(X^2) - (E(X))^2$.

`Démonstration` dans le cas $X(\Omega) = \mathbb{N}$, en notant $m = E(X)$.
Variance : série $\sum\limits_{k\geq 0} (k - m)^2 P(X = k)$ et Moment d'ordre 2 : série $\sum\limits_{k\geq 0} k^2P(X = k)$.
Or les termes sont tous positifs, et $(k - m)^2P(X = k) \underset{k\to+\infty}{\sim} k^2P(X = k)$, donc on peut conclure (critère par équivalence) que ces deux séries ont même nature. Puis dans le cas de convergence :
$(X - E(X))^2 = X^2 - 2XE(X) + E(X)^2$ d'où par linéarité de l'espérance,
$V(X) = E[X^2 - 2XE(X) + E(X)^2] = E(X^2) - 2E(X)E(X) + E(X)^2 = E(X^2) - E(X)^2$ □

Point méthode : pour répondre à la question "X admet-elle une variance ? Si oui la calculer".
Lorsque X admet une espérance (sinon la réponse est non), étudier si X admet un moment d'ordre 2, et le cas échéant utiliser la formule $V(X) = E(X^2) - E(X)^2$ pour la calculer.
En général, pour montrer que le moment d'ordre 2 existe, on étudiera directement la série associée à $E(X^2)$ mais dans certains cas il sera plus judicieux de montrer que $X(X - 1)$ admet une espérance (à l'aide de la formule de transfert). Puis de conclure à l'aide de la linéarité puisque en cas de convergence, $E(X^2) = E(X(X - 1)) + E(X)$.

Propriétés :
Si X admet une variance, alors pour tout $(a, b) \in \mathbb{R}^2$, $aX + b$ admet une variance et $V(aX + b) = a^2V(X)$.

DÉFINITION

Soit X une variable discrète sur $(\Omega, \mathcal{A}, P)$.

1. On dit que X est réduite si $\sigma(X) = 1 = V(X)$
2. Si $\sigma(X) \neq 0$ alors la variable $Y = \dfrac{X - E(X)}{\sigma(X)}$ est appelée variable centrée réduite associée à X.
   En effet, Y vérifie $E(Y) = 0$ et $V(Y) = 1$.

Cas particulier :
Si $V(X) = 0$ alors X est une variable aléatoire constante sur Ω presque sûrement (et réciproquement).

3.3 Moments d'ordre r

DÉFINITION

Soit $r \in \mathbb{N}^*$. Sous réserve d'existence, on appelle moment d'ordre r de X le réel $m_r(X) = E(X^r)$.

Remarque

1. Toute variable aléatoire finie admet des moments à tout ordre, car les sommes finies n'ont pas de problème de convergence !
2. Plus un moment est d'ordre grand, plus la contrainte sur la variable aléatoire est grande. Autrement dit, si X admet un moment d'ordre r alors elle admet un moment d'ordre s pour tout $s \leq r$.

4 Lois usuelles discrètes infinies

Bien sûr toutes les lois usuelles discrètes finies se généralisent telles quelles si Ω est un ensemble infini muni d'une tribu quelconque $\mathcal{A}$. Elles ne sont pas rappelées ici.

4.1 Loi géométrique

DÉFINITION
On effectue une infinité d'épreuves mutuellement indépendantes. Chaque épreuve a deux issues : le succès de probabilité $p \in]0,1[$ et l'échec de probabilité $(1-p)$, et on introduit la variable X égale au rang d'apparition du premier succès, càd au nombre d'épreuves nécessaires à l'obtention du premier succès.
On dit alors que X suit la loi géométrique de paramètre p et on note : $X \hookrightarrow \mathcal{G}(p)$.
$\longrightarrow$ *X est le temps d'attente du premier succès dans un processus de Bernoulli sans mémoire*

`Théorème` Soit $p \in]0,1[$. Si une variable aléatoire X suit la loi géométrique de paramètre p alors
$\boxed{X(\Omega) = \mathbb{N}^*, \text{ et } \forall n \in \mathbb{N}^*, \ P(X=n) = p(1-p)^{n-1}}$

`Démonstration`
Le premier succès peut arriver dès la première épreuve, donc $X(\Omega) = \mathbb{N}^*$.
Introduisons pour tout $i \in \mathbb{N}^*$, les événements S_i (resp. E_i) " la i^e épreuve est un succès (resp. échec)".
Alors $\forall k \in \mathbb{N}^*$, $(X = k) = E_1 \cap E_2 \cap \ldots E_{k-1} \cap S_k$ et par indépendance mutuelle des épreuves
$P(X = k) = P(E_1)P(E_2)...P(E_{k-1})P(S_k) = (1-p)^{k-1}p$.
On vérifie que $\sum\limits_{k=1}^{+\infty} p(1-p)^{k-1} = p\sum\limits_{j=0}^{+\infty}(1-p)^j = p\frac{1}{1-(1-p)} = 1$. (ce qui permet aussi de justifier a posteriori, que presque sûrement on obtient au moins un succès donc que $X(\Omega) = \mathbb{N}^*$) □

Remarque Soit le modèle suivant :
On effectue une succession d'épreuves mutuellement indépendantes, jusqu'à l'obtention du premier succès. Chaque épreuve a deux issues : le succès de probabilité $p \in]0,1[$ et l'échec de probabilité $(1-p)$.
X est la variable aléatoire égale au rang d'apparition du premier succès, c'est-à-dire au nombre d'épreuves effectuées. On pourrait démontrer de la même manière (en utilisant la formule des probabilités composées au lieu de la définition de la mutuelle indépendance) que dans ce cas, on a aussi $X \hookrightarrow \mathcal{G}(p)$.

Espérance et Variance
Si $X \hookrightarrow \mathcal{G}(p)$ alors X admet une espérance et une variance et $\boxed{E(X) = \frac{1}{p}, \ V(X) = \frac{1-p}{p^2}}$.

`Démonstration`
Espérance : étude de la convergence absolue de la série $\sum\limits_{k\geq 1} kp(1-p)^{k-1}$, ce qui revient à la convergence puisque la série est à termes positifs. Or $kp(1-p)^{k-1} = p[k(1-p)^{k-1}]$; on reconnaît le terme général d'une série géométrique dérivée convergente puisque $0 < p < 1 \Rightarrow 0 < 1-p < 1$.
Donc la série converge, donc X admet une espérance et $E(X) = p\sum\limits_{k=1}^{+\infty} k(1-p)^{k-1} = \frac{p}{(1-(1-p))^2} = \frac{1}{p}$.
Moment d'ordre 2 : étude de la convergence de la série $\sum\limits_{k\geq 1} k^2p(1-p)^{k-1}$.
Or pour tout $k \geq 1$, $k^2p(1-p)^{k-1} = p(1-p)[k(k-1)(1-p)^{k-2}] + p[k(1-p)^{k-1}]$, combinaison linéaire de termes généraux de séries géométriques convergentes.
Donc X admet un moment d'ordre 2 et $E(X^2) = p(1-p)\sum\limits_{k=1}^{+\infty} k(k-1)(1-p)^{k-2} + p\sum\limits_{k=1}^{+\infty} k(1-p)^{k-1}$
$= \frac{2p(1-p)}{(1-(1-p))^3} + \frac{p}{(1-(1-p))^2} = \frac{2(1-p)}{p^2} + \frac{1}{p}$ (on pourrait raccourcir en réalisant que la 2e série est la série de l'espérance !)
Donc X admet une variance et $V(X) = E(X^2) - (E(X))^2 = \dfrac{2-p}{p^2} - (\dfrac{1}{p})^2 = \dfrac{1-p}{p^2}$.
OU passer par $E(X(X-1))$ (méthode présentée pour la loi de Poisson) □

4.2 Loi de Poisson et modélisation

Soit $\lambda > 0$. On dit qu'une variable aléatoire X suit une loi de Poisson de paramètre λ, notée $\boxed{\mathcal{P}(\lambda)}$

si $\boxed{X(\Omega) = \mathbb{N} \text{ et } \forall n \in \mathbb{N},\ P(X = n) = \dfrac{\lambda^n}{n!}e^{-\lambda}}$

Remarque

On définit bien ainsi une loi de probabilité car pour tout $n \in \mathbb{N}$, $\frac{\lambda^n}{n!}e^{-\lambda} \geq 0$, et $\sum\limits_{n=0}^{+\infty} \dfrac{\lambda^n}{n!}e^{-\lambda} = e^{-\lambda}e^{\lambda} = 1$.

Espérance et Variance

Si $X \hookrightarrow \mathcal{P}(\lambda)$ alors X admet une espérance et une variance et $\boxed{E(X) = V(X) = \lambda}$.

`Démonstration`

Espérance : étude de la convergence absolue de la série $\sum\limits_{k\geq 0} k\frac{\lambda^k}{k!}e^{-\lambda}$, ce qui revient à la convergence puisque la série est à termes positifs. Or pour $k \geq 1$, $k\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda}[\frac{\lambda^j}{j!}]$ (en posant $j = k-1$) ; on reconnaît le terme général d'une série exponentielle convergente.

Donc X admet une espérance et $E(X) = 0 + \sum\limits_{k=1}^{+\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum\limits_{j=0}^{+\infty}\frac{\lambda^j}{j!} = \lambda e^{-\lambda}e^{\lambda} = \lambda$.

Moment d'ordre 2 : étude de la convergence de la série $\sum\limits_{k\geq 0} k^2\frac{\lambda^k}{k!}e^{-\lambda}$.

Or pour $k \geq 2$, $k^2\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\lambda^k[\frac{k(k-1)}{k!} + \frac{k}{k!}] = e^{-\lambda}\lambda^2[\frac{\lambda^{k-2}}{(k-2)!}] + e^{-\lambda}[\frac{k\lambda^k}{k!}]$ combinaison linéaire d'un terme général de série exponentielle et du terme général de la série de l'espérance, toutes deux convergentes.

Donc X admet un moment d'ordre 2 et $E(X^2) = \sum\limits_{k=0}^{+\infty} e^{-\lambda}\lambda^k\frac{k(k-1)}{k!} + E(X)$

$= 0 + 0 + e^{-\lambda}\lambda^2\sum\limits_{k=2}^{+\infty}[\frac{\lambda^{k-2}}{(k-2)!}] + E(X) = e^{-\lambda}\lambda^2\sum\limits_{j=0}^{+\infty}\frac{\lambda^j}{j!} + \lambda = \lambda^2 + \lambda$.

Donc X admet une variance et $V(X) = E(X^2) - (E(X))^2 = \lambda$.

Variante pour le moment d'ordre 2 :

Montrons que $X(X-1)$ admet une espérance : d'après le théorème de transfert, il faut étudier la convergence absolue de la série $\sum\limits_{k\geq 0} k(k-1)\frac{\lambda^k}{k!}e^{-\lambda}$ ce qui revient à la convergence.

Or pour tout $k \geq 2$, $k(k-1)\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\lambda^2\frac{\lambda^{k-2}}{(k-2)!} = e^{-\lambda}\lambda^2\frac{\lambda^j}{j!}$ (en posant $j = k-2$), et on reconnaît le terme général d'une série exponentielle donc convergente.

Donc $X(X-1)$ admet une espérance et $E(X(X-1)) = 0 + 0 + \sum\limits_{k=2}^{+\infty} k(k-1)\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\lambda^2\sum\limits_{j=0}^{+\infty}\frac{\lambda^j}{j!} = \lambda^2$.

Puis par linéarité de l'espérance (le résultat sous cette forme sera vu en 2e année),
X^2 admet une espérance et $E(X^2) = E(X(X-1)) + E(X) = \lambda^2 + \lambda$.
Donc X admet une variance et $V(X) = E(X^2) - (E(X))^2 = \lambda$. $\square$

bonus pour les plus rapides : Montrer que pour tout entier $r \geq 2$, la variable $X(X-1)...(X-r+1)$ admet une espérance et $E(X(X-1)...(X-r+1)) = \lambda^r$ (moments factoriels de X)

Modélisation : Pourquoi introduire cette loi ?
Imaginons que l'on veuille modéliser le nombre X d'accidents de voiture par jour survenus sur une portion de route bien délimitée. Pour simplifier le modèle, on suppose que chaque voiture qui passe a indépendamment des autres, la même probabilité p d'avoir un accident sur cette portion de route.
Hypothèse 1 : si l'on sait de plus qu'il passe exactement 1000 voitures par jour sur cette portion de route, alors la loi de X est toute trouvée : $X \hookrightarrow \mathcal{B}(1000, p)$.
Mais cette hypothèse 1 pose problème : comment imaginer qu'il y ait un nombre fixe par jour de voitures qui passe par cette portion de route ? (semaine/week-end/vacances ...) C'est dans ce cadre là que la loi de Poisson est utilisée, car cette loi ne demande pas de "valeur maximale" (elle peut théoriquement prendre toutes les valeurs). On pourrait toutefois contrargumenter en disant que permettre un nombre infiniment grand de voitures qui passent sur une portion de route, et donc d'accidents n'est pas non plus réaliste.

En fait, même si théoriquement une variable de loi de Poisson prend toutes les valeurs entières, quand ces valeurs entières sont trop grandes, la probabilité de ces valeurs devient extrêmement faible et donc concrètement, elle n'en prend qu'un petit nombre, même si elle peut prendre exceptionnellement une valeur plus grande.

Exemple : $\lambda = 1$.

valeurs prises	0	1	2	3	> 4
probabilité	$1/e \approx 0.37$	$1/e \approx 0.37$	≈ 0.18	≈ 0.06	≈ 0.02

Donc, une variable de Poisson de paramètre 1, prend concrètement moins de 10 valeurs.