Révisions 5

Correction du sujet ESSEC

ESSEC II 2013

1. (a) S_n est une somme de n variables de Bernouilli indépendantes et de même paramètre p_n, donc elle suit la loi binomiale de de paramètres n et p_n :

$$S_n \hookrightarrow \mathcal{B}(n, p_n)$$

(b) i. On a : $P(S_n = k) = \binom{n}{k}(p_n)^k(1-p_n)^{n-k}$.

ii. Par hypothèses $np_n \xrightarrow[n\to+\infty]{} \lambda$, alors par quotient de limites

$$p_n = \frac{np_n}{n} \xrightarrow[n\to+\infty]{} 0$$

La variable apparaissant en puissance dans $(1-p_n)^n$, on écrit : $(1-p_n)^n = e^{n\ln(1-p_n)}$.
Or $p_n \xrightarrow[n\to+\infty]{} 0$ donc $n\ln(1-p_n) \sim -np_n \xrightarrow[n\to+\infty]{} -\lambda$.
Par continuité de l'exponentielle :

$$(1-p_n)^n \xrightarrow[n\to+\infty]{} e^{-\lambda}$$

iii. On a vu que pour tout $n \geq k$ (suffisant ici puisque n tend vers $+\infty$ et k est fixé) :

$$P(S_n = k) = \binom{n}{k}(p_n)^k(1-p_n)^{n-k} = \frac{1}{(1-p_n)^k} \times \frac{n!}{k!(n-k)!} \times p_n^k \times (1-p_n)^n.$$

On connaît les limites de $\frac{1}{(1-p_n)^k} \xrightarrow[n\to+\infty]{} \frac{1}{(1-0)^k} = 1$ et $(1-p_n)^n \xrightarrow[n\to+\infty]{} e^{-\lambda}$ et $k!$ est une constante, on considère les autres facteurs :

$$\frac{n!}{(n-k)!} \times p_n^k = n(n-1) \times \cdots \times (n-(k-1))p_n^k.$$

Or on sait que $np_n \xrightarrow[n\to+\infty]{} \lambda$, on va donc le faire apparaître :

$$\frac{n!}{(n-k)!} \times p_n^k = \frac{n(n-1) \times \cdots \times (n-(k-1))}{n^k} \times (np_n)^k.$$

Ici on a fait apparaître $(np_n)^k \xrightarrow[n\to+\infty]{} \lambda^k$, il reste à considérer : (et d'après le résultat de l'énoncé, on doit trouver une limite égale à 1 ce qu'on va s'efforcer d'obtenir) :

$$\frac{n(n-1) \times (n-(k-1))}{n^k} = \frac{n}{n} \times \frac{n-1}{n} \times \cdots \times \frac{n-(k-1)}{n}$$

et on a un produit de k facteurs qui tendent tous vers 1 donc sa limite est $1^k = 1$.
Enfin en rassemblant tous nos résultats on a :

$$P(S_n = k) = \frac{n-1}{n} \times \cdots \times \frac{n-(k-1)}{n} \times (np_n)^k \times (1-p_n)^n \times \frac{1}{k!} \times \frac{1}{(1-p_n)^k} \xrightarrow[n\to+\infty]{} \frac{\lambda^k e^{-\lambda}}{k!}$$

par produit de limites.

(c) i. Chaque X_i peut prendre les valeurs 0 ou 1, donc leur maximum peut prendre ces deux valeurs : $N_n(\Omega) = \{0; 1\}$.

ii. Calculons déjà $P(N_n = 0)$: pour avoir $(N_n = 0)$, il faut et il suffit que chacune des variables X_i soit égale à 0 donc :

$$(N_n = 0) = \bigcap_{k=1}^{n} (X_k = 0)$$

et par indépendance des X_i :

$$P(N_n = 0) = P\left(\bigcap_{k=1}^{n} (X_k = 0)\right) = \prod_{k=1}^{n} P(X_k = 0) = \prod_{k=1}^{n} (1 - p_n) = (1 - p_n)^n$$

On a vu que cette quantité avait pour limite $e^{-\lambda}$, on en déduit :

$$\lim_{n \to +\infty} P(N_n = 0) = e^{-\lambda}.$$

iii. On en déduit que

$$P(N_n = 1) = 1 - P(N_n = 0) \xrightarrow[n \to +\infty]{} 1 - e^{-\lambda}$$

puis que la suite (N_n) converge en loi vers une variable aléatoire N qui suit la loi de Bernoulli de paramètre $\left(1 - e^{-\lambda}\right)$.

2. (a) On pose la fonction $f : u \mapsto 1 - u - e^{-u}$ sur $\mathbb{R}_+$, qui est dérivable comme somme et composée de fonctions usuelles dérivables et :

$$f'(u) = -1 + e^{-u} \leq 0 \quad \text{sur } \mathbb{R}_+ \quad \text{car } -u \leq 0 \Rightarrow e^{-u} \leq 1 \text{ par croissance de l'exponentielle}$$

f est donc décroissante sur $\mathbb{R}_+$ et atteint un maximum en $u = 0$, qui vaut $f(0) = 1 - 0 - e^0 = 0$ et :

$$\forall u \in \mathbb{R}_+, \quad f(u) \leq f(0) = 0 \Leftrightarrow \forall u \in \mathbb{R}_+, \quad 1 - u \leq e^{-u}$$

(b) U suit la loi de Bernoulli de paramètre $\frac{\lambda}{n}$, on remplace donc par les valeurs :

$$\sum_{k=0}^{+\infty} \left| P(U = k) - \left(\tfrac{\lambda}{n}\right)^k \tfrac{1}{k!} e^{-\frac{\lambda}{n}} \right| = \left| P(U = 0) - \left(\tfrac{\lambda}{n}\right)^0 \tfrac{1}{0!} e^{-\frac{\lambda}{n}} \right| + \left| P(U = 1) - \left(\tfrac{\lambda}{n}\right)^1 \tfrac{1}{1!} e^{-\frac{\lambda}{n}} \right| + \sum_{k=2}^{+\infty} \left| P(U = k) - \left(\tfrac{\lambda}{n}\right)^k \tfrac{1}{k!} e^{-\frac{\lambda}{n}} \right|$$

donc

$$\sum_{k=0}^{+\infty} \left| P(U = k) - \left(\frac{\lambda}{n}\right)^k \frac{1}{k!} e^{-\frac{\lambda}{n}} \right| = \left| 1 - \frac{\lambda}{n} - e^{-\frac{\lambda}{n}} \right| + \left| \frac{\lambda}{n} - \frac{\lambda}{n} e^{-\frac{\lambda}{n}} \right| + \sum_{k=2}^{+\infty} \left| -\left(\frac{\lambda}{n}\right)^k \frac{1}{k!} e^{-\frac{\lambda}{n}} \right|$$

On a vu que pour tout $u \geq 0$ (ici $u = \frac{\lambda}{n}$), $1 - u - e^{-u} \leq 0$ donc la première valeur absolue vaut l'opposé du nombre à l'intérieur.
D'autre part la deuxième vaut $\frac{\lambda}{n} \left| 1 - e^{-\frac{\lambda}{n}} \right|$ et comme $1 - e^{-u} = -f'(u) \geq 0$ sur $\mathbb{R}_+$ alors toujours en $u = \frac{\lambda}{n}$ la valeur absolue vaut ce nombre.
Enfin à l'intérieur de la somme tous les nombres sont négatifs donc leur valeur absolue est encore une fois leur opposé.
Donc :

$$\sum_{k=0}^{+\infty} \left| P(U = k) - \left(\frac{\lambda}{n}\right)^k \frac{1}{k!} e^{-\frac{\lambda}{n}} \right| = -1 + \frac{\lambda}{n} + e^{-\frac{\lambda}{n}} + \frac{\lambda}{n} - \frac{\lambda}{n} e^{-\frac{\lambda}{n}} + \sum_{k=2}^{+\infty} \left(\frac{\lambda}{n}\right)^k \frac{1}{k!} e^{-\frac{\lambda}{n}}$$

Dans la dernière somme, on sort le facteur constant $e^{-\frac{\lambda}{n}}$ puis on reconnaît une série exponentielle dont on rajoute les deux premiers termes manquants :

$$\sum_{k=0}^{+\infty}\left|P(U=k)-\left(\frac{\lambda}{n}\right)^k\frac{1}{k!}e^{-\frac{\lambda}{n}}\right| = -1+2\frac{\lambda}{n}+e^{-\frac{\lambda}{n}}-\frac{\lambda}{n}e^{-\frac{\lambda}{n}}+e^{-\frac{\lambda}{n}}\left(\sum_{k=0}^{+\infty}\frac{\left(\frac{\lambda}{n}\right)^k}{k!}-1-\frac{\lambda}{n}\right).$$

Enfin cette série exponentielle est convergente et sa somme vaut $e^{\frac{\lambda}{n}}$, puis on simplifie :

$$\begin{aligned}\sum_{k=0}^{+\infty}\left|P(U=k)-\left(\frac{\lambda}{n}\right)^k\frac{1}{k!}e^{-\frac{\lambda}{n}}\right| &= -1+2\frac{\lambda}{n}+e^{-\frac{\lambda}{n}}-\frac{\lambda}{n}e^{-\frac{\lambda}{n}}+e^{-\frac{\lambda}{n}}\left(e^{\frac{\lambda}{n}}-1-\frac{\lambda}{n}\right)\\ &= -1+2\frac{\lambda}{n}+e^{-\frac{\lambda}{n}}-\frac{\lambda}{n}e^{-\frac{\lambda}{n}}+1-e^{-\frac{\lambda}{n}}-\frac{\lambda}{n}e^{-\frac{\lambda}{n}}\end{aligned}$$

On élimine les termes qui se compensent et on factorise par $2\frac{\lambda}{n}$:

$$\sum_{k=0}^{\infty}\left|P(U=k)-\left(\frac{\lambda}{n}\right)^k\frac{1}{k!}e^{-\frac{\lambda}{n}}\right| = \frac{2\lambda}{n}\left[1-e^{-\frac{\lambda}{n}}\right]$$

(c) On applique à nouveau la question (a) : on isole d'abord $1-e^{-u}$ dans l'inégalité du (a) :

$$1-u\leq e^{-u} \Longleftrightarrow 1-e^{-u}\leq u$$

qu'on applique à $u=\frac{\lambda}{n}$, ce qui donne :

$$1-e^{-\frac{\lambda}{n}}\leq\frac{\lambda}{n}$$

et enfin on multiplie par $\frac{2\lambda}{n}\geq 0$:

$$\sum_{k=0}^{+\infty}\left|P(U=k)-\left(\frac{\lambda}{n}\right)^k\frac{1}{k!}e^{-\frac{\lambda}{n}}\right|\leq 2\left(\frac{\lambda}{n}\right)^2.$$

3. (a) L'inégalité triangulaire donne pour tous a et b réels :

$$|a-b|=|a+(-b)|\leq|a|+|(-b)|=|a|+|b|$$

On l'applique ici avec $a=P(U=k-i)$ et $b=P(V=k-i)$ et comme une probabilité est toujours positive on obtient :

$$0\leq|P(U=k-i)-P(V=k-i)\ |\leq P(U=k-i)+P(V=k-i)$$

On veut prouver la convergence d'une série sans obtenir sa somme, on passe par un théorème de comparaison et les hypothèses obtenues nous font utiliser une majoration.

Les séries de termes généraux $P(U=k-i)$ et $P(V=k-i)$ convergent car une fois enlevé les termes négatifs (qui valent 0), on a la somme des probabilités de chacune de ces variables aléatoires, qui converge vers 1.
On obtient donc la convergence de la série majorante (et même sa somme égale à 2, ce qui ne nous intéresse pas) donc par théorème de comparaison des séries à termes positifs, la série de terme général $|P(U=k-i)-P(V=k-i)|$ converge.
On note A_i sa somme : $A_i=\sum_{k=0}^{+\infty}|P(U=k-i)-P(V=k-i)|$.

On remarque la ressemblance avec le résultat de la question 2)c), on va donc utiliser celui-ci. On commence par ré-indexer la somme avec $k' = k - i$:

$$A_i = \sum_{k=-i}^{+\infty} |P(U=k) - P(V=k)|.$$

et comme les variables U et V ont des probabilités nulles d'être négatives,

$$A_i = \sum_{k=0}^{+\infty} |P(U=k) - P(V=k)|.$$

Enfin, comme dans la question 2, U suit une loi de Bernoulli de paramètre $\frac{\lambda}{n}$, et V suit la loi de Poisson de paramètre $\frac{\lambda}{n}$, on remplace par la valeur de ses probabilités :

$$A_i = \sum_{k=0}^{+\infty} \left| P(U=k) - e^{-\frac{\lambda}{n}} \frac{\left(\frac{\lambda}{n}\right)^k}{k!} \right|$$

qui est exactement la série qui a été majorée à la question 2. Donc on obtient bien :

$$A_i \leq 2\left(\frac{\lambda}{n}\right)^2$$

(b) A nouveau on utilise un théorème de comparaison, en utilisant la série convergente liée à la variable Z (attention à ne pas vouloir utiliser une série de Riemann, car c'est i et pas n qui varie).
Comme $0 < \lambda \leq n$,

$$A_i \times P(Z=i) \leq 2\left(\frac{\lambda}{n}\right)^2 \times P(Z=i)$$

avec $\sum\limits_{i=0}^{+\infty} 2\left(\frac{\lambda}{n}\right)^2 P(Z=i)$ qui converge vers $2\left(\frac{\lambda}{n}\right)^2$ (voir question précédente).
Dans les deux cas, le théorème de comparaison des séries à termes positifs permet de conclure à la convergence de la série de terme général $A_i \times P(Z=i)$.

(c) Cette fois-ci k est fixé et c'est i qui tend vers $+\infty$: alors pour tout $i \geq k+1$, les probabilités $P(U=k-i)$ et $P(V=k-i)$ sont nulles donc

$$\forall i \geq k+1, \quad |P(U=k-i) - P(V=k-i)| \times P(Z=i) = 0$$

et la série converge (c'est une somme finie).

(d) i. On a vu que tous les termes pour $i \geq k+1$ sont nuls, on sépare alors les termes $i = k-1$ et $i = k$ qu'on calcule à part :

$$|P(U=0) - P(V=0)| \times P(Z=k) = \left|1 - \frac{\lambda}{n} - e^{-\frac{\lambda}{n}}\right| P(Z=k)$$

et

$$|P(U=1) - P(V=1)\ | \times P(Z=k-1) = \left|\frac{\lambda}{n} - \lambda e^{-\frac{\lambda}{n}}\right| P(Z=k-1)$$

donc

$$\begin{aligned} B_k &= \sum_{i=0}^{k} |P(U=k-i) - P(V=k-i)| \times P(Z=i) \\ &= \left|1 - \frac{\lambda}{n} - e^{-\frac{\lambda}{n}}\right| P(Z=k) + \left|\frac{\lambda}{n} - \lambda e^{-\frac{\lambda}{n}}\right| P(Z=k-1) \\ &\quad + \sum_{i=0}^{k-2} |P(U=k-i) - P(V=k-i)| \times P(Z=i). \end{aligned}$$

Enfin, pour $i \leq k-2$, $k-i \geq 2$ donc $P(U=k-i)=0$ (rappelons que $U(\Omega)=\{0;1\}$). Donc $\forall k \in [\![0;k-2]\!]$,

$$|P(U=k-i)-P(V=k-i)|\times P(Z=i) = |-P(V=k-i)|\times P(Z=i) = P(V=k-i)\times P(Z=i)$$

car une probabilité est toujours positive. Enfin on obtient :

$$B_k = \left|1-\frac{\lambda}{n}-e^{-\frac{\lambda}{n}}\right| P(Z=k) + \left|\frac{\lambda}{n}-\frac{\lambda}{n}e^{-\frac{\lambda}{n}}\right| P(Z=k-1) + \sum_{i=0}^{k-2} P(V=k-i)P(Z=i)$$

ii. On reconnaît la loi d'une somme de variables aléatoires, on va alors utiliser les probabilités totales avec le système complet d'évènement $(Z=i)_{i\geq 0}$:

$$P(V+Z=k) = \sum_{i=0}^{+\infty} P\big([V+Z=k]\cap[Z=i]\big) = \sum_{i=0}^{+\infty} P\big([V=k-i]\cap[Z=i]\big)$$

De plus les variables aléatoires sont indépendantes, et on va séparer les termes cherchés :

$$\begin{aligned} P(V+Z=k) &= \sum_{i=0}^{+\infty} P(V=k-i)P(Z=i) \\ &= \sum_{i=0}^{k-2} P(V=k-i)P(Z=i) + \sum_{i=k-1}^{+\infty} P(V=k-i)P(Z=i). \end{aligned}$$

Enfin cette dernière somme est positive ou nulle comme somme de termes positifs ou nuls donc :

$$P(V+Z=k) \geq \sum_{i=0}^{k-2} P(V=k-i)P(Z=i).$$

iii. On obtient la majoration suivante de B_k :

$$B_k \leq \left|1-\frac{\lambda}{n}-e^{-\frac{\lambda}{n}}\right| P(Z=k) + \left|\frac{\lambda}{n}-\frac{\lambda}{n}e^{-\frac{\lambda}{n}}\right| P(Z=k-1) + P(V+Z=k)$$

Remarquons que tous les facteurs devant les probabilités sont constants (seul k est variable ici) donc on obtient une combinaison linéaire de trois séries convergentes car chacune est la somme des probabilités d'une certaine variable aléatoire (Z, Z puis $V+Z$).
Par théorème de comparaison des séries à termes positifs (B_k est la somme d'une série à termes positifs donc c'est bien un nombre positif), la série de terme général B_k converge.

4. (a) On procède comme précédemment ; on utilise l'inégalité triangulaire pour majorer le terme général, et on reconnaît des séries convergentes :

$$|P(Z+U=k)-P(Z+V=k)\ | \leq P(Z+U=k)+P(Z+V=k)$$

qui est la somme des termes généraux de deux séries convergentes, toujours pour les mêmes raisons (sommes des probabilités des variables aléatoires $U+Z$ et $V+Z$).
Par théorème de comparaison des séries à termes positifs, la série de terme général $|P(Z+U=k)-P(Z+V=k)\ |$ est convergente.

(b) On applique les probabilités totales avec le système complet d'évènement $(Z=i)_{i\geq 0}$:

$$\begin{aligned} &|P(Z+U=k)-P(Z+V=k)| \\ =\ &\left|\sum_{i=0}^{+\infty} P\big([Z+U=k]\cap[Z=i]\big) - \sum_{i=0}^{+\infty} P\big([Z+V=k]\cap[Z=i]\big)\right| \end{aligned}$$

On rassemble les deux séries convergentes et on simplifie les probabilités :

$$\begin{aligned}&|P(Z+U=k)-P(Z+V=k)|\\ =\ &\left|\sum_{i=0}^{+\infty}\left[P\big([U=k-i]\cap[Z=i]\big)-P\big([V=k-i]\cap[Z=i]\big)\right]\right|.\end{aligned}$$

Les variables U et Z d'une part et V et Z d'autre part sont indépendantes, on obtient :

$$|P(Z+U=k)-P(Z+V=k)|=\left|\sum_{i=0}^{+\infty}[P(U=k-i)P(Z=i)-P(V=k-i)P(Z=i)]\right|.$$

Enfin, en factorisant par $P(Z=i)$ et par inégalité triangulaire on obtient :

$$|P(Z+U=k)-P(Z+V=k)|\leq\sum_{i=0}^{+\infty}|P(U=k-i)-P(V=k-i)|\times P(Z=i).$$

Enfin on somme pour k allant de 0 à $+\infty$ (les deux séries sont bien convergentes donc on peut le faire) :

$$\sum_{k=0}^{+\infty}|P(Z+U=k)-P(Z+V=k)|\leq\sum_{k=0}^{+\infty}\sum_{i=0}^{+\infty}|P(U=k-i)-P(V=k-i)|\times P(Z=i).$$

et enfin puisqu'on a admis que l'échange de l'ordre des sommes était licite, on obtient bien le résultat demandé.
Enfin on remarque que cette somme s'écrit encore :

$$\sum_{k=0}^{+\infty}|P(Z+U=k)-P(Z+V=k)|\leq\sum_{i=0}^{+\infty}A_iP(Z=i).$$

et on a vu que $A_i\leq 2\left(\frac{\lambda}{n}\right)^2$ donc en multipliant par $P(Z=i)$ et en sommant pour i allant de 0 à $+\infty$ avec deux séries qui convergent bien :

$$\sum_{k=0}^{+\infty}|P(Z+U=k)-P(Z+V=k)|\leq\sum_{k=0}^{+\infty}A_iP(Z=i)\leq\sum_{k=0}^{+\infty}2\left(\frac{\lambda}{n}\right)^2P(Z=i)=2\left(\frac{\lambda}{n}\right)^2$$

puisque $Z(\Omega)\subset\mathbb{N}$ donc $\sum\limits_{i=0}^{+\infty}P(Z=i)=1$.

5. (a) On va utiliser astucieusement l'inégalité triangulaire. On fait d'abord apparaître tous les nouveaux termes à l'intérieur de la valeur absolue :

$$\begin{gathered}|P(U_1+\cdots+U_n=k)-P(V_1+\cdots+V_n=k)|=\\ |P(U_1+\cdots+U_n=k)-P(U_1+\cdots+U_{n-1}+V_n=k)+\\ P(U_1+\cdots+U_{n-1}+V_n=k)-P(U_1+\cdots+U_{n-2}+V_{n-1}+V_n=k)+\\ \cdots+P(U_1+V_2+\cdots+V_n=k)-P(V_1+\cdots+V_n=k)|\end{gathered}$$

et l'inégalité triangulaire donne alors :

$$\begin{gathered}|P(U_1+\cdots+U_n=k)-P(V_1+\cdots+V_n=k)\ |\leq\\ |P(U_1+\cdots+U_n=k)-P(U_1+\cdots+U_{n-1}+V_n=k)\ |+\\ |P(U_1+\cdots+U_{n-1}+V_n=k)-P(U_1+\cdots+U_{n-2}+V_{n-1}+V_n=k)\ |+\\ \cdots+|P(U_1+V_2+\cdots+V_n=k)-P(V_1+\cdots+V_n=k)|\end{gathered}$$

(b) On en déduit immédiatement (sous réserve de convergence des séries du membre de droite) que :

$$\sum_{k=0}^{+\infty} |P(U_1 + \cdots + U_n = k) - P(V_1 + \cdots + V_n = k)| \leq$$

$$\sum_{k=0}^{+\infty} |P(U_1 + \cdots + U_n = k) - P(U_1 + \cdots + U_{n-1} + V_n = k)| +$$

$$\sum_{k=0}^{+\infty} |P(U_1 + \cdots + U_{n-1} + V_n = k) - P(U_1 + \cdots + U_{n-2} + V_{n-1} + V_n = k)\ | +$$

$$\cdots + \sum_{k=0}^{+\infty} |P(U_1 + V_2 + \cdots + V_n = k) - P(V_1 + \cdots + V_n = k)|$$

On remarque qu'à chaque fois une seule variable aléatoire diffère (U_n et V_n dans la première somme, U_{n-1} et V_{n-1} dans la seconde, etc...).

On pose alors à chaque fois Z_j la variable restante (égale des deux côtés) : $Z_n = U_1 + \cdots + U_{n-1}$ dans la première somme, $Z_{n-1} = U_1 + \cdots + U_{n-2} + V_n$ dans la seconde, etc... jusqu'à $Z_1 = V_2 + \cdots + V_n$ dans la dernière somme (la n-ème). Remarquer qu'on a numéroté Z avec le même numéro que la variable qui change.

$$\sum_{k=0}^{+\infty} |P(U_1 + \cdots + U_n = k) - P(V_1 + \cdots + V_n = k)|$$
$$\leq \sum_{j=1}^{n} \left[\sum_{k=0}^{+\infty} |P(U_j + Z_j = k) - P(V_j + Z_j = k)| \right]$$

On reconnaît alors dans chacune de ses sommes le résultat de la question 4 (toutes les hypothèses sont vérifiées avec $Z_j(\Omega) \subset \mathbb{N}$, $U_j \hookrightarrow \mathcal{B}\left(\frac{\lambda}{n}\right)$ et $V_j \hookrightarrow \mathcal{P}\left(\frac{\lambda}{n}\right)$ et ces trois variables indépendantes car Z_j ne dépend que des variables d'indice autre que j, qui sont indépendantes de U_j et V_j).

On en déduit que ces sommes convergent bien (ce qui justifie l'écriture de départ) puis en appliquant à chacune la question 4 :

$$\sum_{k=0}^{+\infty} |P(U_1 + \cdots + U_n = k) - P(V_1 + \cdots + V_n = k)\ | \leq \sum_{j=1}^{n} \frac{2\lambda^2}{n^2} = \frac{2\lambda^2}{n}$$

(c) La somme $\sum\limits_{j=1}^{n} U_j$ suit la loi binomiale de paramètres n et $\frac{\lambda}{n}$ comme somme de variables de Bernoulli indépendantes et de même paramètre.

De même la somme $\sum\limits_{j=1}^{n} V_j$ suit la loi de Poisson de paramètres $n \times \frac{\lambda}{n} = \lambda$ comme somme de variables indépendantes suivant des lois de Poisson.

On a alors $P(X = k) = P(U_1 + \cdots + U_n = k)$ pour tout k, et $P(V_1 + \cdots + V_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}$ pour tout k donc en remplaçant dans l'inégalité de 5)b) on obtient bien :

$$\sum_{k=0}^{+\infty} \left| P(X = k) - e^{-\lambda}\frac{\lambda^k}{k!} \right| \leq \frac{2\lambda^2}{n}$$

6. (a) X_i suit la loi de Bernoulli de paramètre $\frac{\Delta V}{V}$ (la probabilité que la bactérie soit dans le prélèvement).

(b) On a clairement $N = \sum_{i=1}^{n} X_i$ et comme ces variables de Bernoulli sont indépendantes et de même paramètre, N suit la loi binomiale de paramètres n et $\frac{\Delta V}{V}$.

7. On remarque tout d'abord que pour tout entier k on peut obtenir F comme suit :

- Si $k < 0$, $F(k) = 0$.
- Si $k = 0$, $F(k) = P(U = 0) = e^{-\lambda}$.
- Si $k \geq 1$, $F(k) = F(k-1) + P(U = k-1)$.

On va donc calculer $F(k)$ par récurrence, et de plus la probabilité $P(U = k) = e^{-\lambda}\frac{\lambda^k}{k!}$ vérifie :

$$P(U = k) = \frac{\lambda}{k} P(U = k-1)$$

ce qui va permettre de calculer $P(U = k)$ par récurrence également.

Pour minimiser le nombre d'opérations on va calculer ces termes récurrents en même temps.

Enfin pour un x réel, on remarque que $F(x) = F\lfloor x \rfloor$ ce qui permet de se ramener à un entier.

On peut alors écrire :

```
def fct_rep_Poisson(x,lambda):
  if x < 0:
    F = 0
  else:
    n = np.floor(x)
    p = np.exp(-lambda)
    F = 0
    for k in range(0,n+1):
        F = F+p
        p = lambda/(k+1)*p
    return(F)
```

8. (a) On applique la question 5.(c) de la partie I avec N qui suit une loi binomiale de paramètres n et $\frac{\lambda}{n}$ et posant $\lambda = \frac{n\Delta V}{V} = c\Delta V$, puis U qui suit la loi de Poisson de paramètre $\lambda = c\Delta V$, et N et U qui sont indépendantes.

On obtient :

$$\sum_{k=0}^{+\infty} |P(N = k) - P(U = k)| \leq \frac{2\lambda^2}{n} = \frac{2c^2(\Delta V)^2}{n}.$$

On en déduit (tous les termes de la somme sont positifs) que :

$$\sum_{k=0}^{K} |P(N = k) - P(U = k)| \leq \sum_{k=0}^{+\infty} |P(N = k) - P(U = k)| \leq \frac{2c^2(\Delta V)^2}{n}$$

Or par inégalité triangulaire la somme de gauche vérifie :

$$\sum_{k=0}^{K} |P(N = k) - P(U = k)| \geq \left|\sum_{k=0}^{K} P(N = k) - P(U = k)\right| = |P(N \leq K) - P(U \leq K)|$$

En rassemblant et en remarquant que $\frac{c}{n} = \frac{1}{V}$ on obtient :

$$|P(N \leq K) - P(U \leq K)| \leq \frac{2c^2(\Delta V)^2}{n} = \frac{2c(\Delta V)^2}{V}.$$

(b) Dans cette application numérique, on obtient en posant E l'erreur commise :

$$E = |P(N \leq K) - P(U \leq K)| \leq \frac{2c(\Delta V)^2}{V} = \frac{2c(10^{-3})^2}{10^3} = \frac{2c}{10^9}$$

(c) On a toujours :

$$E \leq \frac{2c(\Delta V)^2}{V} \leq \frac{2 \times 10^6(10^{-3})^2}{V} = \frac{2}{V}$$

donc pour garantir $E \leq 10^{-6}$ il faut :

$$E \leq \frac{2}{V} \leq 10^{-6} \Leftrightarrow V \geq 2 \times 10^6$$

c'est-à-dire qu'il faut un volume supérieur à 2 000 000 m^3.

(d) En effectuant les mêmes opérations que précédemment on obtient :

$$|P(N \leq K) - P(U \leq K)| \leq \frac{2(c\Delta V)\min(c\Delta V, 2)}{n}$$

puis on remplace $\frac{c}{n} = \frac{1}{V}$ et on remarque que $\min(c\Delta V, 2) \leq 2$ donc on a :

$$|P(N \leq K) - P(U \leq K)| \leq \frac{2\Delta V \min(c\Delta V, 2)}{V} \leq 4\frac{\Delta V}{V}$$

(e) On obtient

$$E \leq 4\frac{\Delta V}{V} = \frac{4 \times 10^{-3}}{V}$$

donc pour garantir que $E \leq 10^{-6}$ il faut que :

$$\frac{4 \times 10^{-3}}{V} \leq 10^{-6} \Leftrightarrow V \geq 4 \times 10^3$$

donc il faut un volume supérieur à 4000 m^3.

9. Sur cet intervalle, h est dérivable comme somme, produit et composée de fonctions dérivables et on a :

$$h'(x) = 1 \times \ln(1+x) + (1+x) \times \frac{1}{1+x} - 1 = \ln(1+x) + 1 - 1 = \ln(1+x)$$

et on résout avec l'exponentielle strictement croissante :

$$\ln(1+x) > 0 \Leftrightarrow 1 + x > e^0 = 1 \Leftrightarrow x > 0$$

qui va donner le tableau suivant :

x	-1		0		$+\infty$
$h'(x)$	‖	$-$	0	$+$	‖
$h(x)$	‖	$\searrow$	0	$\nearrow$	‖
$h(x)$	‖	$+$	0	$+$	‖

10. On écrit la définition de $E(Y)$ et on sépare en deux sommes selon que les valeurs soient avant ou après α :

$$E(Y) = \sum_{y\in Y(\Omega)} yP(Y=y) = \sum_{y\in Y(\Omega), y<\alpha} yP(Y=y) + \sum_{y\in Y(\Omega), y\geq\alpha} yP(Y=y)$$

La première somme est positive car Y est à valeurs positives et les probabilités sont toujours positives.

Dans la deuxième, on a $y \geq \alpha$ donc on peut majorer :

$$E(Y) \geq 0 + \sum_{y\in Y(\Omega), y\geq\alpha} \alpha P(Y=y) = \alpha \sum_{y\in Y(\Omega), y\geq\alpha} P(Y=y) = \alpha P(Y\geq\alpha)$$

et on obtient bien :

$$E(Y) \geq \alpha P(Y\geq\alpha) \Leftrightarrow P(Y\geq\alpha) \leq \frac{E(Y)}{\alpha}$$

11. (a) D'après le théorème de transfert on a sous réserve de convergence :

$$E(e^{uX}) = \sum_{k=0}^{+\infty} e^{uk} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{+\infty}\frac{(\lambda e^u)^k}{k!}$$

On reconnaît une série exponentielle qui converge, e^{uX} admet donc une espérance et :

$$E(e^{uX}) = e^{-\lambda}\times e^{\lambda e^u} = e^{\lambda(e^u-1)}$$

(b) On calcule le côté gauche :

$$E\left(e^{u_0(X-(1+\varepsilon)\lambda)}\right) = E\left(e^{u_0X}\times e^{-u_0\lambda(1+\varepsilon)}\right)$$

et comme $e^{-u_0\lambda(1+\varepsilon)}$ est une constante la linéarité de l'espérance donne :

$$E\left(e^{u_0(X-(1+\varepsilon)\lambda)}\right) = e^{-u_0\lambda(1+\varepsilon)}E(e^{u_0X}) = e^{-u_0\lambda(1+\varepsilon)}\times e^{\lambda(e^{u_0}-1)} = e^{\lambda(e^{u_0}-1-u_0-u_0\varepsilon)}$$

Il faut donc obtenir

$$h(\varepsilon) = (1+\varepsilon)\ln(1+\varepsilon)-\varepsilon = 1+u_0+u_0\varepsilon-e^{u_0} \Longleftrightarrow (1+\varepsilon)\ln(1+\varepsilon) = 1+\varepsilon+u_0(1+\varepsilon)-e^{u_0}$$

$$\Longleftrightarrow (1+\varepsilon)\left[\ln(1+\varepsilon)-1\right] = u_0(1+\varepsilon)-e^{u_0}$$

On étudie alors la fonction $f(u) = u(1+\varepsilon)-e^u-(1+\varepsilon)\left[\ln(1+\varepsilon)-1\right]$ qui est dérivable sur $\mathbb{R}_+$ et vérifie :

$$f'(u) = (1+\varepsilon)-e^u$$

qui est décroissante et s'annule en $u=\ln(1+\varepsilon)$ donc f' est positive puis négative.

Enfin f est croissante puis décroissante et admet un maximum en $u_1=\ln(1+\varepsilon)$, qui vaut :

$$\begin{aligned} f(u_1) &= (1+\varepsilon)\ln(1+\varepsilon)-e^{\ln(1+\varepsilon)}-(1+\varepsilon)\left[\ln(1+\varepsilon)-1\right] \\ &= (1+\varepsilon)\ln(1+\varepsilon)-(1+\varepsilon)-(1+\varepsilon)\left[\ln(1+\varepsilon)-1\right] = 0. \end{aligned}$$

On en déduit que f s'annule au point $u_0=u_1=\ln(1+\varepsilon)$, et en ce point on a bien :

$$E(e^{u_0(X-(1+\varepsilon)\lambda)}) = e^{-\lambda h(\varepsilon)}$$

(c) i. Pour tout $u>0$,

$$P(\lambda^{-1}(X-\lambda)\geq\varepsilon) = P(X-\lambda\geq\lambda\varepsilon) = P(u(X-\lambda)\geq u\lambda\varepsilon) = P\left(e^{u(X-\lambda)}\geq e^{\lambda\varepsilon u}\right)$$

ii. La variable aléatoire $e^{u(X-\lambda)}$ est positive et admet une espérance, on lui appliqué l'inégalité de Markov :

$$P\left(e^{u(X-\lambda)} \geq e^{\lambda\varepsilon u}\right) \leq \frac{E\left(e^{u(X-\lambda)}\right)}{e^{\lambda\varepsilon u}} = E\left(e^{u(X-\lambda)-\lambda\varepsilon u}\right) = E\left(e^{u(X-\lambda(1+\varepsilon))}\right)$$

donc en reprenant la question précédente on obtient :

$$\forall u > 0,\ P(\lambda^{-1}(X-\lambda) \geq \varepsilon) \leq E\left(e^{u(X-\lambda(1+\varepsilon))}\right).$$

En particulier, c'est vrai pour $u = u_0$ qui est bien strictement positif donc :

$$P(\lambda^{-1}(X-\lambda) \geq \varepsilon) \leq E\left(e^{u_0(X-\lambda(1+\varepsilon))}\right) = e^{-\lambda h(\varepsilon)}$$

(d) - Premier cas : $\varepsilon \geq 1$, on a alors $-\varepsilon \leq -1$ donc :

$$h(-\varepsilon) = 0 \quad \text{donc} \quad e^{-\lambda h(-\varepsilon)} = e^0 = 1$$

et comme une probabilité est toujours inférieure ou égale à 1, l'inégalité est bien vérifiée.

- Deuxième cas : $0 < \varepsilon < 1$: On reprend les questions précédentes, avec quelques changements :
Comme il faut changer le sens de l'inégalité pour utiliser Markov, on utilise cette fois-ci $u < 0$: Pour tout $u > 0$,

$$P(\lambda^{-1}(X-\lambda) \leq -\varepsilon) = P(X-\lambda \leq -\varepsilon\lambda) = P(u(X-\lambda) \geq -u\varepsilon\lambda) = P\left(e^{u(X-\lambda)} \geq e^{-u\varepsilon\lambda}\right)$$

On applique à nouveau l'inégalité de Markov et on obtient :

$$P(\lambda^{-1}(X-\lambda) \leq -\varepsilon) = P\left(e^{u(X-\lambda)} \geq e^{-u\varepsilon\lambda}\right) \leq \frac{E\left(e^{u(X-\lambda)}\right)}{e^{-u\varepsilon\lambda}} = E\left(e^{u(X-\lambda)+u\varepsilon\lambda}\right)$$

On reprend la question 11.(b), et on montre qu'il existe $u_1 < 0$ tel que :

$$E\left(e^{u_1(X-\lambda)+u_1\varepsilon\lambda}\right) = e^{-\lambda h(-\varepsilon)}$$

On se ramène à nouveau à une équation sur u, puis on étudie une fonction et on obtient une unique solution :

$$u_1 = \ln(1-\varepsilon)$$

qui est bien strictement négatif.
L'inégalité précédente étant vraie pour tout $u < 0$, elle est vraie en u_1 et enfin on en déduit :

$$P(\lambda^{-1}(X-\lambda) \leq -\varepsilon) \leq e^{-\lambda h(-\varepsilon)}$$

On obtient bien pour tout $\varepsilon > 0$:

$$P(\lambda^{-1}(X-\lambda) \leq -\varepsilon) \leq e^{-\lambda h(-\varepsilon)}$$

12. (a) On transforme :

$$P(X \leq \alpha) = P(X-\lambda \leq \alpha-\lambda) = P\big(\lambda^{-1}(X-\lambda) \leq \lambda^{-1}(\alpha-\lambda)\big)$$

puis on applique la question 11.(d) pour $\varepsilon = \lambda^{-1}(\lambda-\alpha)$ qui est bien strictement positif car $\lambda > 2000 \geq \alpha$. On obtient :

$$P(X \leq \alpha) \leq e^{-\lambda h(\lambda^{-1}(\alpha-\lambda))} = e^{-\lambda h(\lambda^{-1}\alpha-1)}.$$

(b) On construit patiemment l'inégalité : on a $\lambda > 2000$ donc :

$$\lambda^{-1} < \frac{1}{2000} \quad \text{puis} \quad \lambda^{-1}\alpha < \frac{\alpha}{2000} \quad \text{puis} \quad \lambda^{-1}\alpha - 1 < \frac{\alpha}{2000} - 1$$

Ces deux réels sont compris entre 0 et 1 (on reconnaît le $-\varepsilon$ de la question précédente à gauche, négatif, et plus grand que -1 car $\lambda^{-1}\alpha > 0$), et à droite $\alpha \leq 2000$ donc $\frac{\alpha}{2000} \leq 1$. La fonction h est de plus décroissante sur $]-1;0]$ donc :

$$h\left(\lambda^{-1}\alpha - 1\right) > h\left(\frac{\alpha}{2000} - 1\right) \quad \text{puis} \quad -\lambda h\left(\lambda^{-1}\alpha - 1\right) < -\lambda h\left(\frac{\alpha}{2000} - 1\right)$$

et enfin par stricte croissance de l'exponentielle :

$$e^{-\lambda h\left(\lambda^{-1}\alpha - 1\right)} < e^{-\lambda h\left(\frac{\alpha}{2000} - 1\right)}$$

On reprend alors l'inégalité sur $P(X - \alpha)$:

$$P(X - \alpha) \leq e^{-\lambda h\left(\lambda^{-1}\alpha - 1\right)} \leq e^{-\lambda h\left(\frac{\alpha}{2000} - 1\right)}$$

(c) Cette équation est équivalente à :

$$2000 \times h\left(\frac{x}{2000} - 1\right) = \ln(100) \Leftrightarrow h\left(\frac{x}{2000} - 1\right) = \frac{\ln(100)}{2000}$$

On reprend alors l'étude de h et on rajoute les limites (évidente en $+\infty$, avec le changement de variable $y = 1 + x$ en -1) au tableau de variation :

x	-1		0		$+\infty$
$h'(x)$	‖	$-$	0	$+$	‖
$h(x)$	1	$\searrow$	0	$\nearrow$	$+\infty$
$h(x)$	‖	$+$	0	$+$	‖

Or

$$\frac{\ln(100)}{2000} = \frac{\ln(10^2)}{2000} = \frac{2\ln(10)}{2000} = \frac{\ln(10)}{1000} < 1$$

donc l'équation $h(y) = \frac{\ln(100)}{2000}$ admet une unique solution $y \in]-1;0]$, et une unique solution $z \in [0;+\infty[$.
De plus, on a $x \in [0, 2000]$ donc $\dfrac{x}{2000} - 1 \in [-1, 0]$ donc

$$\frac{x}{2000} - 1 = y \Leftrightarrow x = 2000(y+1)$$

donc l'équation initiale admet deux solutions :

$$\alpha_0 = 2000(y+1) \in [0;2000] \quad \text{car} \quad y \in [-1;0] \quad \text{et} \quad \alpha_1 \in [2000;+\infty[\quad \text{car} \quad z \geq 0.$$

Elle admet une unique solution α_0 comprise entre 0 et 2000.

(d) On admet que $1865 < \alpha_0$.
Par croissance de la fonction de répartition, on en déduit si $\lambda > 2000$:

$$P(X < 1865) \leq P(X \leq 1865) \leq P(X \leq \alpha_0) \leq e^{-2000h\left(\frac{\alpha_0}{2000}-1\right)}$$

Or α_0 étant solution de l'équation précédente, on a

$$2000h\left(\frac{\alpha_0}{2000} - 1\right) = \ln(100)$$

donc

$$P(X < 1865 \leq e^{-2000h\left(\frac{\alpha_0}{2000}-1\right)} = e^{-\ln(100)} = \frac{1}{100}$$

(e) Rappelons qu'on approxime la loi binomiale du nombre de bactéries dans le prélèvement par la loi de Poisson que suit X.
Avec 1600 bactéries dans le prélèvement, on obtient $X = 1600 < 1865$.
Or on a vu précédemment que $P_{(\lambda > 2000)}(X < 1865) \leq \frac{1}{100}$.
Ceci signifie qu'on a un risque très faible d'avoir $\lambda > 2000$ (d'autant que 1600 est largement plus petit que 1865, la probabilité est probablement bien plus faible en réalité).
Or la tolérance est fixée à 2000 bactéries par litre, donc 2000×10^3 bactéries par mètre cube, et enfin $\lambda = c\Delta V = 2000 \times 10^3 \times 10^{-3} = 2000$.
On voit donc que le risque que le seuil de tolérance soit dépassé est extrêmement faible, et on peut autoriser le bassin sans prendre de risque (en fait on peut minimiser encore plus ce risque en faisant plusieurs prélèvements et en faisant la synthèse des résultats, c'est la théorie de l'estimation par intervalle de confiance).