Correction CB 2 - Maths II ECS 2022

Questions préliminaires

1. (a) Le résultat s'obtient par récurrence immédiate. En effet, on a :

$$C_0 = 0$$

qui vérifie donc la relation.
Et pour $n \in \mathbb{N}$, on suppose que $C_n = X_1 + \cdots + X_n$. On a alors :

$$C_{n+1} = C_n + X_{n+1} = X_1 + \cdots + X_n + X_{n+1}.$$

Donc on a pour tout $n \in \mathbb{N}$:

$$\boxed{C_n = X_1 + \cdots + X_n.}$$

(b) C'est une question de modélisation.
X_i est le nombre d'achats supplémentaires nécessaires pour obtenir une nouvelle vignette d'un type différent. C'est donc le numéro du premier succès dans une expérience de Bernoulli répétée où le succès est d'obtenir une nouvelle vignette. C'est bien une loi géométrique.
De plus, pour X_i, on a déjà $i-1$ vignettes et donc il reste $n-(i-1)$ vignettes différentes à récupérer. Par équiprobabilité, la probabilité du succès est $\frac{n-i+1}{n}$ et donc :

$$\boxed{X_i \hookrightarrow \mathcal{G}\left(\frac{n-i+1}{n}\right).}$$

Partie I

2. (a) Pour $k \geqslant 2$, on a :

$$\frac{1}{k} - \frac{1}{k+1} = \frac{k+1-k}{k(k+1)} = \frac{1}{k^2+k}.$$

Comme $k^2 + k \geqslant k^2 > 0$, par décroissance de la fonction inverse sur $\mathbb{R}_+^\star$, on a :

$$\boxed{\frac{1}{k} - \frac{1}{k+1} \leqslant \frac{1}{k^2}.}$$

On a de même :

$$\boxed{\frac{1}{k^2} \leqslant \frac{1}{k-1} - \frac{1}{k}.}$$

(b) Pour tout $k \geqslant 2$, on a :

$$\frac{1}{k} - \frac{1}{k+1} \leqslant \frac{1}{k^2}.$$

En sommant de 2 à n, on a :

$$\underbrace{\sum_{k=2}^{n}\left(\frac{1}{k} - \frac{1}{k+1}\right)}_{=\frac{1}{2}-\frac{1}{n+1}} \leqslant \underbrace{\sum_{k=2}^{n}\frac{1}{k^2}}_{=S_n-1}.$$

Et en réordonnant les termes, on a :

$$\boxed{\frac{3}{2} - \frac{1}{n+1} \leqslant S_n.}$$

On a de manière similaire en sommant l'autre inégalité :

$$\boxed{S_n \leqslant 2 - \frac{1}{n}.}$$

(c) (S_n) est clairement croissante (somme partielle d'une série à termes positifs). On a d'après l'inégalité précédente :

$$S_n \leqslant 2 - \frac{1}{n} \leqslant 2.$$

Et donc (S_n) est majorée.

$\boxed{(S_n) \text{ converge donc}}$ d'après le théorème de la limite monotone. Notons S sa limite.

Puis, en passant à la limite l'encadrement précédent, on :

$$\boxed{\frac{3}{2} \leqslant S \leqslant 2.}$$

(d) On reprend la première inégalité vérifiée pour tout $k \geqslant 2$:

$$\frac{1}{k} - \frac{1}{k+1} \leqslant \frac{1}{k^2} \leqslant \frac{1}{k-1} - \frac{1}{k}.$$

On sait que les séries de termes généraux $\frac{1}{k} - \frac{1}{k+1}$, $\frac{1}{k^2}$ et $\frac{1}{k-1} - \frac{1}{k}$ convergent. On peut donc regarder les restes de ces séries. On somme de $n+1$ à $+\infty$ l'inégalité précédente, et on obtient :

$$\sum_{k=n+1}^{+\infty} \left(\frac{1}{k} - \frac{1}{k+1}\right) \leqslant \sum_{k=n+1}^{+\infty} \frac{1}{k^2} \leqslant \sum_{k=n+1}^{+\infty} \left(\frac{1}{k-1} - \frac{1}{k}\right).$$

On a pour $N > n$:

$$\sum_{k=n+1}^{N} \left(\frac{1}{k} - \frac{1}{k+1}\right) = \frac{1}{n+1} - \frac{1}{N+1} \xrightarrow[N \to +\infty]{} \frac{1}{n+1}.$$

Donc :

$$\sum_{k=n+1}^{+\infty} \left(\frac{1}{k} - \frac{1}{k+1}\right) = \frac{1}{n+1}.$$

On a de même :

$$\sum_{k=n+1}^{+\infty} \left(\frac{1}{k-1} - \frac{1}{k}\right) = \frac{1}{n}.$$

Puis, on a :

$$\sum_{k=n+1}^{+\infty} \frac{1}{k^2} = \sum_{k=1}^{+\infty} \frac{1}{k^2} - \sum_{k=1}^{n} \frac{1}{k^2} = S - S_n.$$

Donc :

$$\boxed{\frac{1}{n+1} \leqslant S - S_n \leqslant \frac{1}{n}.}$$

(e)

```
def S_approx():
    n = 10**7
    S = 0
    for i in range(1,n):
        S = S + 1/(i*i)
    return S
```

3. (a) ln est $\mathcal{C}^2$ sur $\mathbb{R}_+^\star$ et on a pour $x \in \mathbb{R}_+^\star$:

$$\ln'(x) = \frac{1}{x} \quad \text{et} \quad \ln''(x) = -\frac{1}{x^2} < 0.$$

Donc ln est concave. En particulier, ln est sous ses tangentes. On a en particulier avec la tangente en 1 :

$$\forall x \in \mathbb{R}, \ \ln(x) \leqslant x - 1.$$

Pour $x = \frac{n+1}{n}$, on obtient :

$$\boxed{\forall n \geqslant 1,\ \ln\left(\frac{n+1}{n}\right) \leqslant \frac{1}{n}.}$$

Et pour $x = \frac{n}{n+1}$, on obtient :

$$\forall n \geqslant 1,\ \ln\left(\frac{n}{n+1}\right) \leqslant \frac{n}{n+1} - 1 = -\frac{1}{n+1}.$$

Et en multipliant par -1, on a :

$$\boxed{\forall n \geqslant 1,\ \frac{1}{n+1} \leqslant \ln\left(\frac{n+1}{n}\right).}$$

(b) Sommons les inégalités précédentes de 1 à $n-1$. On obtient :

$$\sum_{k=1}^{n-1} \frac{1}{k+1} \leqslant \sum_{k=1}^{n-1} \ln\left(\frac{k+1}{k}\right) \leqslant \sum_{k=1}^{n-1} \frac{1}{k}.$$

Or : $\sum_{k=1}^{n-1} \frac{1}{k+1} = \sum_{k=2}^{n} \frac{1}{k} = H_n - 1$. Et de même : $\sum_{k=1}^{n-1} \frac{1}{k} = H_n - \frac{1}{n}$.
On a aussi :

$$\sum_{k=1}^{n-1} \ln\left(\frac{k+1}{k}\right) = \sum_{k=1}^{n-1} (\ln(k+1) - \ln(k)) = \ln(n) - \ln(1) = \ln(n).$$

Donc :

$$H_n - 1 \leqslant \ln(n) \leqslant H_n - \frac{1}{n}.$$

On en déduit :

$$\boxed{\underbrace{\frac{1}{n}}_{\geqslant 0} \leqslant \underbrace{H_n - \ln(n)}_{=u_n} \leqslant 1.}$$

(c) On a pour tout $n \geqslant 1$:

$$\begin{aligned} u_{n+1} - u_n &= H_{n+1} - \ln(n+1) - H_n + \ln(n) \\ &= \frac{1}{n+1} - \ln\left(\frac{n+1}{n}\right) \leqslant 0 \end{aligned}$$

où le signe est obtenu à partir d'une question précédente. Donc (u_n) est décroissante. (u_n) est également minorée.
Donc $\boxed{(u_n) \text{ converge.}}$ Notons γ sa limite. En passant à la limite l'inégalité $0 \leqslant u_n \leqslant 1$, on a :

$$\boxed{0 \leqslant \gamma \leqslant 1.}$$

4. Comme $X_i \hookrightarrow \mathcal{G}\left(\frac{n-i+1}{n}\right)$, X_i admet une espérance. Puis par linéarité de l'espérance, $\boxed{C_n \text{ admet une espérance.}}$
On a de plus :

$$\begin{aligned} E(C_n) &= E(X_1) + \cdots + E(X_n) = \frac{n}{n-1+1} + \frac{n}{n-2+1} + \cdots + \frac{n}{n-n+1} \\ &= n\left(\sum_{k=1}^{n} \frac{1}{n-k+1}\right) \underset{k'=n-k+1}{=} n\left(\sum_{k'=1}^{n} \frac{1}{k'}\right) = \boxed{nH_n.} \end{aligned}$$

5. De même, les X_i admettent tous une variance, donc $\boxed{C_n \text{ admet une variance}}$ comme somme.

De plus, comme les X_i sont indépendants, on a :

$$\begin{aligned}
\mathrm{Var}(C_n) &= \sum_{i=1}^{n} \mathrm{Var}(X_i) = \sum_{i=1}^{n} \frac{1-\frac{n-i+1}{n}}{\left(\frac{n-i+1}{n}\right)^2} \\
&= \sum_{i=1}^{n} \frac{n^2}{(n-i+1)^2} \times \frac{i-1}{n} = \sum_{i=1}^{n} \frac{n(i-1)}{(n-i+1)^2} \\
&= n\sum_{i=1}^{n} \frac{i-1}{(n-i+1)^2} = n\sum_{i=1}^{n} \left(\frac{i-1-n}{(n-i+1)^2} + \frac{n}{(n-i+1)^2}\right) \\
&= n\sum_{i=1}^{n} \frac{-1}{n-i+1} + n^2\sum_{i=1}^{n} \frac{1}{(n-i+1)^2} \underset{k=n-i+1}{=} -n\sum_{k=1}^{n} \frac{1}{k} + n^2\sum_{k=1}^{n} \frac{1}{k^2} \\
&= \boxed{n^2 S_n - nH_n.}
\end{aligned}$$

6. (a) C_n admet une espérance et une variance. Donc l'inégalité de Bienaymé-Tchebychev s'applique. On a :

$$\forall \epsilon > 0,\ \mathbb{P}(|C_n - E(C_n)| \geqslant \epsilon) \leqslant \frac{\mathrm{Var}(C_n)}{\epsilon^2}.$$

Commençons par remarquer que $E(C_n) = nH_n$ et $\mathrm{Var}(C_n) = n^2 S_n - nH_n$. Ainsi pour $\epsilon = an$, on obtient :

$$\mathbb{P}(|C_n - nH_n| \geqslant an) \leqslant \frac{n^2 S_n - nH_n}{(an)^2}.$$

Or $H_n \geqslant 0$, donc :

$$\frac{n^2 S_n - nH_n}{(an)^2} = \frac{S_n}{a^2} - \frac{H_n}{a^2 n} \leqslant \frac{S_n}{a^2}.$$

De plus (S_n) est croissante et donc inférieur à sa limite S. On en déduit que :

$$\boxed{\mathbb{P}(|C_n - nH_n| \geqslant an) \leqslant \frac{S}{a^2}.}$$

(b) Soit $c > 1$. Posons $a = c - 1$. On a bien $a > 0$. Donc :

$$\mathbb{P}(|C_n - nH_n| \geqslant (c-1)n) \leqslant \frac{S}{(c-1)^2}.$$

Or, on a :

$$\begin{aligned}
\mathbb{P}(|C_n - nH_n| \geqslant (c-1)n) &= \mathbb{P}(|C_n - n(u_n + \ln(n))| \geqslant (c-1)n) \\
&= \mathbb{P}\left(\left|\frac{C_n}{n} - (u_n + \ln(n))\right| \geqslant (c-1)\right).
\end{aligned}$$

Or :

$$\left|\frac{C_n}{n} - (u_n + \ln(n))\right| \geqslant \left|\frac{C_n}{n} - \ln(n)\right| - |u_n| \geqslant \left|\frac{C_n}{n} - \ln(n)\right| - 1.$$

Donc :

$$\left[\left|\frac{C_n}{n} - (u_n + \ln(n))\right| \geqslant (c-1)\right] \supset \left[\left|\frac{C_n}{n} - \ln(n)\right| - 1 \geqslant (c-1)\right]$$

et ainsi :

$$\mathbb{P}\left(\left|\frac{C_n}{n} - (u_n + \ln(n))\right| \geqslant (c-1)\right) \geqslant \mathbb{P}\left(\left|\frac{C_n}{n} - \ln(n)\right| - 1 \geqslant (c-1)\right).$$

Et enfin :

$$\boxed{\mathbb{P}\left(\left|\frac{C_n}{n} - \ln(n)\right| \geqslant c\right) \leqslant \frac{S}{(c-1)^2}.}$$

(c) On a montré que $S \leqslant 2$. Donc :

$$\mathbb{P}\left(\left|\frac{C_n}{n} - \ln(n)\right| \geqslant c\right) \leqslant \frac{2}{(c-1)^2}.$$

Pour $c = 6$, on obtient :

$$\mathbb{P}\left(\frac{C_n}{n} \in]\ln(n) - 6, \ln(n) + 6[\right) \geqslant 1 - \frac{2}{(6-1)^2}.$$

Puis comme $\ln(n) \approx 13,816$, on a :

$$\left[\frac{C_n}{n} \in]\ln(n) - 6, \ln(n) + 6[\right] \subset \left[\frac{C_n}{n} \in [13,81 - 6, 13,82 + 6]\right].$$

On a ainsi :

$$\boxed{\mathbb{P}\left(\frac{C_n}{n} \in [7,81, 19,82]\right) \geqslant \frac{23}{25} = 0,92.}$$

Le sujet semble contenir une faute puisque l'énoncé a $19,92$ à la place de $19,82$. Qu'importe avec une seconde inclusion d'événements, on a :

$$\boxed{\mathbb{P}\left(\frac{C_n}{n} \in [7,81, 19,92]\right) \geqslant \mathbb{P}\left(\frac{C_n}{n} \in [7,81, 19,82]\right) \geqslant 0,92.}$$

Partie II

7. Comme $\max : \mathbb{R}^n \to \mathbb{R}$ est $\mathcal{C}^0$ sur $\mathbb{R}^n$, M_n est bien une variable aléatoire.
Notons F_{M_n} sa fonction de répartition.
Soit $x \in \mathbb{R}$. On a :

$$\begin{aligned} F_{M_n}(x) &= \mathbb{P}(M_n \leqslant x) = \mathbb{P}(\max(T_1, \ldots, T_n) \leqslant x) \\ &= \mathbb{P}\left(\bigcap_{i=1}^{n} [T_i \leqslant x]\right) = \prod_{i=1}^{n} \mathbb{P}(T_i \leqslant x) \text{ (indépendance)} \\ &= \Big(\mathbb{P}(T_1 \leqslant x)\Big)^n = \left\{\begin{array}{ll} 0 & \text{si } x < 0 \\ (1 - \mathrm{e}^{-x})^n & \text{si } x \geqslant 0 \end{array}\right. \text{. (loi exponentielle de paramètre 1)} \\ & \text{(loi identique)} \end{aligned}$$

Clairement F_{M_n} est $\mathcal{C}^1$ sauf éventuellement en 0. De plus, on a :

$$\lim_{x \to 0^+} (1 - \mathrm{e}^{-x})^n = 0$$

et donc F_{M_n} est $\mathcal{C}^0$ sur $\mathbb{R}$.
Donc $\boxed{M_n \text{ est à densité.}}$
De plus, pour $x \in \mathbb{R}^\star$, on a :

$$F'_{M_n}(x) = \left\{\begin{array}{ll} 0 & \text{si } x < 0 \\ n\mathrm{e}^{-x}(1 - \mathrm{e}^{-x})^{n-1} & \text{si } x > 0 \end{array}\right. .$$

Donc la fonction f_n définie sur $\mathbb{R}$ par :

$$f_n(x) = \left\{\begin{array}{ll} 0 & \text{si } x \leqslant 0 \\ n\mathrm{e}^{-x}(1 - \mathrm{e}^{-x})^{n-1} & \text{si } x > 0 \end{array}\right.$$

$\boxed{\text{est une densité de } M_n.}$

8. (a) On a :

$$\boxed{g : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \left\{ \begin{array}{ll} (n+1)\mathrm{e}^{-(n+1)x} & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{array} \right. \end{array} \right. .}$$

Pour tout $x > 0$ et pour tout $t \in]0, x[$, on a donc :

$$\begin{aligned} f_n(t)g(x-t) &= n\mathrm{e}^{-t}(1-\mathrm{e}^{-t})^{n-1} \times (n+1)\mathrm{e}^{-(n+1)(x-t)} \\ &= n\mathrm{e}^{-t}(1-\mathrm{e}^{-t})^{n-1}(n+1)\mathrm{e}^{-(n+1)x}\mathrm{e}^{(n+1)t} \\ &= \boxed{(n+1)\mathrm{e}^{-(n+1)x}n\mathrm{e}^{t}(\mathrm{e}^{t}-1)^{n-1}.} \end{aligned}$$

(b) On remarque de plus que pour $x \leqslant 0$ ou pour $x \geqslant 0$ mais $t \notin]0, x[$, on a :

$$\boxed{f_n(t)g(x-t) = 0.}$$

Montrons maintenant le résultat par récurrence.

- **Initialisation :** pour $n = 1$, on a Z_1 qui suit une loi exponentielle de paramètre 1 et $\boxed{f_1 \text{ est bien une densité associée}}$ (il suffit de vérifier la formule pour $n = 1$).
- **Hérédité :** Soit $n \in \mathbb{N}^\star$. On suppose que f_n est une densité de $\sum_{i=1}^n Z_i$.
  D'après le lemme des coalitions, on a $\sum_{i=1}^n Z_i$ et Z_{n+1} sont indépendantes. Comme une densité de Z_{n+1} est bornée, le produit de convolution converge et est une densité de la somme. Notons h la densité de $\sum_{i=1}^n Z_i + Z_{n+1}$ définie par :

$$h(x) = \int_{-\infty}^{+\infty} f_n(t)g(x-t)\mathrm{d}t.$$

  Avec les calculs précédents, on a pour $x \leqslant 0$, $h(x) = 0$ et pour $x > 0$:

$$\begin{aligned} h(x) &= \int_0^x (n+1)\mathrm{e}^{-(n+1)x}\mathrm{e}^{t}(\mathrm{e}^{t}-1)^{n-1}\mathrm{d}t = (n+1)\mathrm{e}^{-(n+1)x}\int_0^x n\mathrm{e}^{t}(\mathrm{e}^{t}-1)^{n-1}\mathrm{d}t \\ &= (n+1)\mathrm{e}^{-(n+1)x}\left[(\mathrm{e}^{t}-1)^{n}\right]_0^x = (n+1)\mathrm{e}^{-(n+1)x}(\mathrm{e}^{x}-1)^{n} \\ &= (n+1)\mathrm{e}^{-x}(1-\mathrm{e}^{-x})^{n} = \boxed{f_{n+1}(x).} \end{aligned}$$

  Donc $\boxed{f_{n+1} \text{ est une densité de } \sum_{i=1}^{n+1} Z_i.}$

Donc $\boxed{f_n \text{ est bien une densité de } \sum_{i=1}^n Z_i \text{ pour tout } n \in \mathbb{N}^\star.}$

9. f est clairement $\mathcal{C}^0$ par opérations sur les fonctions usuelles et positives (même strictement).
De plus pour $A, B \in \mathbb{R}$, on a :

$$\int_A^B f(x)\mathrm{d}x = \int_A^B \mathrm{e}^{-x}\mathrm{e}^{-\mathrm{e}^{-x}}\mathrm{d}x = \left[\mathrm{e}^{-\mathrm{e}^{-x}}\right]_A^B = \underbrace{\mathrm{e}^{-\mathrm{e}^{-B}}}_{\xrightarrow[B\to+\infty]{}1} - \underbrace{\mathrm{e}^{-\mathrm{e}^{-A}}}_{\xrightarrow[A\to-\infty]{}0}.$$

Donc $\boxed{\int_{-\infty}^{+\infty} f(x)\mathrm{d}x \text{ converger et vaut } 1.}$

Donc $\boxed{f \text{ est une densité de probabilité.}}$

10. (a) Soit $x \in \mathbb{R}$. On a :

$$F_{W_n}(x) = \mathbb{P}(W_n \leqslant x) = \mathbb{P}\left(\sum_{i=1}^n Z_i - \ln(n) \leqslant x\right) = \mathbb{P}\left(\sum_{i=1}^n Z_i \leqslant x + \ln(n)\right).$$

On a montré que $\sum_{i=1}^n Z_i$ et M_n partagent une densité et donc la même fonction de répartition. On a donc :

$$F_{W_n}(x) = F_{M_n}(x+\ln(n)) = \left\{ \begin{array}{ll} \left(1-\mathrm{e}^{-x-\ln(n)}\right)^n & \text{si } x+\ln(n) > 0 \\ 0 & \text{si } x+\ln(n) \leqslant 0 \end{array} \right. = \boxed{\left\{ \begin{array}{ll} \left(1-\frac{\mathrm{e}^{-x}}{n}\right)^n & \text{si } x > -\ln(n) \\ 0 & \text{si } x \leqslant -\ln(n) \end{array} \right.}.$$

(b) Soit $x \in \mathbb{R}$. Pour n suffisamment grand, on a $x > -\ln(n)$. Donc pour n suffisamment grand, on a :

$$F_{W_n}(x) = \left(1 - \frac{\mathrm{e}^{-x}}{n}\right)^n = \exp\left(n \ln\left(1 - \frac{\mathrm{e}^{-x}}{n}\right)\right).$$

Comme $\frac{\mathrm{e}^{-x}}{n} \xrightarrow[n\to+\infty]{} 0$, on a :

$$\ln\left(1 - \frac{\mathrm{e}^{-x}}{n}\right) \underset{n\to+\infty}{\sim} -\frac{\mathrm{e}^{-x}}{n} \quad \text{puis} \quad n\ln\left(1 - \frac{\mathrm{e}^{-x}}{n}\right) \underset{n\to+\infty}{\sim} -\mathrm{e}^{-x}.$$

Donc en passant aux limites, on a : $F_{W_n}(x) \xrightarrow[n\to+\infty]{} \mathrm{e}^{-\mathrm{e}^{-x}}$. Donc en tout point de continuité de G, on a bien $F_{W_n}(x) \xrightarrow[n\to+\infty]{} G(x)$ donc :

$$\boxed{W_n \xrightarrow{\mathcal{L}} W}$$

où $\boxed{W \text{ suit la loi de Gumbel.}}$

11. (a)

```python
def simulV(n):
    C = 0
    for i in range(1,n+1):
        C = C + rd.geometric((n-i+1)/n)
    y = C/n - np.log(n)
    return y
```

(b)

```python
V = np.zeros(1000)
for i in range(1000):
    V[i] = simulV(n)
```

(c) Les diagrammes bâtons sont une représentation des fréquences observées et avec 1000 simulations c'est une bonne approximation de la loi de V_n.
On remarque que plus n est grand plus le diagramme suit la courbe théorique de la loi de Gumbel. On peut donc conjecturer que $\boxed{\text{la suite de variable } (V_n) \text{ converge en loi vers la loi de Gumbel.}}$

12. U est à valeurs positives et on peut facilement vérifier que $\alpha > 0$. Donc V est à valeurs dans $\mathbb{N}^\star$.
Pour tout $k \in \mathbb{N}^\star$, on a :

$$\mathbb{P}(V = k) = \mathbb{P}(\lfloor \alpha U \rfloor + 1 = k) = \mathbb{P}(k - 1 \leqslant \alpha U < k) = \mathbb{P}\left(\frac{k-1}{\alpha} \leqslant U < \frac{k}{\alpha}\right) = F_U\left(\frac{k}{\alpha}\right) - F_U\left(\frac{k-1}{\alpha}\right)$$

où F_U est la fonction de répartition de U de loi $\mathcal{E}(p)$. On a donc puisque $\frac{k}{\alpha} \geqslant 0$ et $\frac{k-1}{\alpha} \geqslant 0$:

$$\begin{aligned}
\mathbb{P}(V = k) &= \left(1 - \mathrm{e}^{p\frac{k}{\alpha}}\right) - \left(1 - \mathrm{e}^{p\frac{k-1}{\alpha}}\right) = \mathrm{e}^{p\frac{k\ln(1-p)}{p}} - \mathrm{e}^{p\frac{(k-1)\ln(1-p)}{p}} \\
&= (1-p)^k - (1-p)^{k-1} = (1-p)^{k-1}(1 - (1-p)) = \boxed{(1-p)^{k-1}p.}
\end{aligned}$$

Donc $\boxed{V \text{ suit bien la loi géométrique de paramètre } p.}$

13. (a) On a :

$$V - \alpha U = \lfloor \alpha U \rfloor - \alpha U + 1.$$

Or pour tout $x \in \mathbb{R}$, on a :

$$\lfloor x \rfloor \leqslant x < \lfloor x \rfloor + 1$$

et donc :

$$0 < \lfloor x \rfloor - x + 1 \leqslant 1.$$

Donc $\boxed{V - \alpha U \text{ est à valeur dans }]0,1] \text{ et donc } \textit{a fortiori} \text{ dans } [0,1].}$

On a donc par croissance de l'espérance :

$$0 \leqslant E(X) \leqslant 1 \quad \text{et} \quad 0 \leqslant E(X^2) \leqslant 1.$$

Puis d'après la formule de Kœnig-Huygens :

$$\boxed{\operatorname{Var}(X) = E(X^2) - E(X)^2 \leqslant 1 - 0^2 = 1.}$$

(b) On a :

$$0 \leqslant (|X| - |Y|)^2 \leqslant X^2 - 2|XY| + Y^2.$$

On a ainsi : $|XY| \leqslant \frac{1}{2}(X^2+Y^2)$. Si X et Y admettent une variance, elles admettent un moment d'ordre 2, puis par domination $E(XY)$ existe.
Ainsi par linéarité :

$$E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$$

existe et $\boxed{\operatorname{Cov}(X, Y) \text{ existe bien.}}$
On calcule ensuite :

$$\begin{aligned} \operatorname{Var}(X + Y) &= \operatorname{Var}(X) + \operatorname{Var}(Y) + 2\operatorname{Cov}(X, Y), \\ \operatorname{Var}(X - Y) &= \operatorname{Var}(X) + \operatorname{Var}(Y) - 2\operatorname{Cov}(X, Y). \end{aligned}$$

En faisant la somme, on obtient : $\operatorname{Var}(X+Y)+\operatorname{Var}(X-Y) = 2(\operatorname{Var}(X)+\operatorname{Var}(Y))$. Comme $\operatorname{Var}(Y) \geqslant 0$, on en déduit :

$$\boxed{\operatorname{Var}(X + Y) \leqslant 2(\operatorname{Var}(X) + \operatorname{Var}(Y)).}$$

(c) On applique le résultat précédent avec $X = V - \alpha U$ et $Y = (\alpha + 1)U$ qui admettent toutes deux des variances. On a :

$$\underbrace{\operatorname{Var}(X + Y)}_{=\operatorname{Var}(U-V)} \leqslant 2(\operatorname{Var}(V - \alpha U) + \operatorname{Var}((\alpha + 1)U)) \leqslant 2 + 2(\alpha + 1)^2\operatorname{Var}(U) \leqslant \boxed{2 + 2 \times \frac{(\alpha + 1)^2}{p^2}.}$$

14. (a) On a :

$$V_n' - W_n' = \frac{X_1' + \cdots + X_n'}{n} - \frac{Y_1 + \cdots + Y_n}{n} = \frac{1}{n}\sum_{i=1}^{n}(X_i' - Y_i).$$

Les variables aléatoires sont mutuellement indépendantes car les variables Y_1, ..., Y_n le sont et chaque X_i' ne dépend que de Y_i. Ainsi :

$$\operatorname{Var}(V_n' - W_n') = \frac{1}{n^2}\sum_{i=1}^{n}\operatorname{Var}(X_i' - Y_i).$$

Maintenant, considérons les différentes valeurs de i :

- Pour $i = 1$, on a $X_1' = 1$ et $Y_1 \hookrightarrow \mathcal{E}\left(\frac{n-1+1}{n}\right)$. Donc :

$$\operatorname{Var}(X_1' - Y_1) = \operatorname{Var}(1 - Y_1) = \operatorname{Var}(Y_1) = \frac{1}{1^2} = 1.$$

- Pour $i \geqslant 2$, on peut appliquer l'inégalité de la question précédente. On a :

$$p = \frac{n - i + 1}{n} \in]0, 1[$$

et :

$$\alpha = -\frac{p}{\ln(1 - p)} = -\frac{\frac{n-i+1}{n}}{\ln\left(1 - \frac{n-i+1}{n}\right)} = -\frac{n - i + 1}{n} \times \frac{1}{\ln\left(\frac{i-1}{n}\right)} = \alpha_i.$$

On obtient donc :

$$\operatorname{Var}(X_i' - Y_i) \leqslant 2 + 2\frac{(1 - \alpha_i)^2}{p^2} = 2 + 2\left(\frac{n}{n - i + 1}\right)^2(1 - \alpha_i)^2.$$

Ainsi :

$$\boxed{\operatorname{Var}(V_n' - W_n') \leqslant \frac{1}{n^2}\sum_{i=2}^{n}\left[1 + \left(\frac{n}{n-i+1}\right)^2 (1-\alpha_i)^2\right].}$$

(b) i. ϕ est continue sur $]0,1[$ par opérations sur les fonctions usuelles.
De plus, en 0^+, on a :

$$\frac{1-z}{\ln(z)} \xrightarrow[z\to 0^+]{} 0$$

et donc :

$$\boxed{\phi(z) \xrightarrow[z\to 0^+]{} 1.}$$

En 1^-, on a :

$$\ln(z) = \ln(1+(z-1)) = (z-1) - \frac{(z-1)^2}{2} + \underset{z\to 1^-}{o}((z-1)^2).$$

On a donc :

$$\begin{aligned}
\phi(z) &= \frac{1}{(1-z)^2}\left(1 + \frac{1-z}{\ln(z)}\right)^2 \\
&= \frac{1}{(1-z)^2}\left(1 + \frac{1-z}{(z-1) - \frac{(z-1)^2}{2} + \underset{z\to 1^-}{o}((z-1)^2)}\right)^2 \\
&= \frac{1}{(1-z)^2}\left(1 + \frac{-1}{1 + \frac{z-1}{2} + \underset{z\to 1^-}{o}(z-1)}\right)^2 \\
&= \frac{1}{(1-z)^2}\left(1 - \left[1 - \frac{z-1}{2} + \underset{z\to 1^+}{o}(z-1)\right]\right)^2 \\
&= \frac{1}{4}\left(1 + \underset{z\to 1^-}{o}(1)\right)^2 \boxed{\xrightarrow[z\to 1^-]{} \frac{1}{4}.}
\end{aligned}$$

Donc $\boxed{\phi \text{ peut être prolongée en une fonction continue sur } [0,1]}$ en posant :

$$\boxed{\phi(0) = 1 \quad \text{et} \quad \phi(1) = \frac{1}{4}.}$$

ii. On a pour $i \geqslant 2$:

$$\begin{aligned}
\left(\frac{n}{n-i+1}\right)^2 (1-\alpha_i)^2 &= \left(\frac{n}{n-i+1}\right)^2 \left(1 - \left[-\frac{n-i+1}{n} \times \frac{1}{\ln\left(\frac{i-1}{n}\right)}\right]\right)^2 \\
&= \frac{1}{\left(1 - \frac{i-1}{n}\right)^2}\left(1 + \left[1 - \frac{i-1}{n}\right] \times \frac{1}{\ln\left(\frac{i-1}{n}\right)}\right)^2 = \Phi\left(\frac{i-1}{n}\right).
\end{aligned}$$

Donc :

$$\frac{1}{n}\sum_{i=2}^{n}\left(\frac{n}{n-i+1}\right)^2 (1-\alpha_i)^2 = \frac{1}{n}\sum_{i=2}^{n}\Phi\left(\frac{i-1}{n}\right).$$

On reconnaît à un terme près une somme de Riemann d'une fonction continue. Détaillons-le. On a :

$$\boxed{\frac{1}{n}\sum_{i=2}^{n}\left(\frac{n}{n-i+1}\right)^2 (1-\alpha_i)^2 = \frac{1}{n}\sum_{i=1}^{n}\Phi\left(\frac{i-1}{n}\right) - \frac{1}{n}\Phi(0) \xrightarrow[n\to+\infty]{} \int_0^1 \Phi(x)\mathrm{d}x - 0.}$$

(c) $V'_n - W'_n$ admet une variance. On peut donc appliquer l'inégalité de Bienaymé-Tchebychev :

$$\forall \epsilon > 0,\ \mathbb{P}(|V'_n - W'_n - E(V'_n - W'_n)| \geqslant \epsilon) \leqslant \frac{\operatorname{Var}(V'_n - W'_n)}{\epsilon^2}.$$

On a par linéarité de l'espérance $E(V'_n - W'_n) = \frac{1}{n}\sum_{i=1}^n (E(X'_i) - E(Y_i))$.
Or on a pour $i \geqslant 2$: $E(Y_i) = \frac{n}{n-i+1}$.
Et on a montré précédemment que $X'_i = \lfloor \alpha_i Y_i \rfloor + 1$ suit la loi $\mathcal{G}\left(\frac{n-i+1}{n}\right)$. Donc : $E(X'_i) = \frac{n}{n-i+1}$.
On vérifie que de même pour $i = 1$, on a bien : $E(X'_1) = E(1) = 1 = E(Y_1)$.
On a donc :

$$\boxed{E(V'_n - W'_n) = 0.}$$

Puis avec les questions précédentes, on a :

$$\begin{aligned}
\operatorname{Var}(V'_n - W'_n) &\leqslant \frac{1}{n^2} + \frac{2}{n^2}\sum_{i=2}^n \left(\frac{n}{n-i+1}\right)^2 (1-\alpha_i)^2 \\
&\leqslant \underbrace{\frac{1}{n^2} + 2\frac{n-1}{n} + \frac{2}{n} \times \underbrace{\frac{1}{n}\sum_{i=2}^n \left(\frac{n}{n-i+1}\right)^2 (1-\alpha_i)^2}_{\xrightarrow[n\to+\infty]{} A}}_{\xrightarrow[n\to+\infty]{} 0}.
\end{aligned}$$

Comme $\operatorname{Var}(V'_n - W'_n) \geqslant 0$, on a par encadrement :

$$\boxed{\operatorname{Var}(V'_n - W'_n) \xrightarrow[n\to+\infty]{} 0.}$$

Puis par encadrement d'une probabilité positive, on a à nouveau :

$$\boxed{\forall \epsilon > 0,\ \mathbb{P}(|V'_n - W'_n| \geqslant \epsilon) \xrightarrow[n\to+\infty]{} 0}$$

c'est-à-dire :

$$\boxed{V'_n - W'_n \xrightarrow{P} 0.}$$

(d) La variable Y_i suit la loi $\mathcal{E}\left(\frac{n-i+1}{n}\right)$ donc $\frac{Y_i}{n}$ suit la loi $\mathcal{E}(n-i+1)$. **Attention !** La loi exponentielle n'est pas stable par transformation affine. En revanche, on peut multiplier par un scalaire.
Ainsi la variable :

$$W'_n = \sum_{i=1}^n \frac{Y_i}{n} - \ln(n)$$

suit la même loi que la variable aléatoire :

$$\sum_{i=1}^n Z_{n-i+1} - \ln(n) = \sum_{j=1}^n Z_j - \ln(n) = W_n.$$

Et donc (W'_n) tend en loi vers la loi de Gumbel.
On écrit alors :

$$V'_n = (V'_n - W'_n) + W'_n.$$

Comme $V'_n - W'_n \xrightarrow{P} 0$, d'après le théorème de Slutsky, on a que (V'_n) tend en loi vers la loi de Gumbel.

(e) On a de la même manière (V_n) qui converge en loi vers la loi de Gumbel car :

$$V_n = \frac{X_1 + \cdots + X_n}{n} - \ln(n)$$

et

$$V'_n = \frac{X'_1 + \cdots + X'_n}{n} - \ln(n)$$

où les X_i ont les mêmes lois que les X_i'. Donc V_n et V_n' ont la même loi.
On peut donc écrire :

$$\forall a, b \in \mathbb{R},\ a < b \Rightarrow \mathbb{P}\left(a \leqslant V_n \leqslant b\right) \xrightarrow[n\to+\infty]{} \mathrm{e}^{-\mathrm{e}^{-b}} - \mathrm{e}^{-\mathrm{e}^{-a}}.$$

Pour $n = 10^6$, on peut considérer que cette limite est une valeur approchée de la probabilité. Ainsi avec $b = 3,20$ et $a = -1,17$, on a :

$$\mathbb{P}\left(-1,17 \leqslant \frac{C_n}{n} - \ln(n) \leqslant 3,20\right) \approx \mathrm{e}^{-\mathrm{e}^{-3,20}} - \mathrm{e}^{-\mathrm{e}^{-1,17}} \approx 0,96 - 0,04 = 0,92.$$

Et en réordonnant dans la probabilité et en utilisant $\ln(n) \approx 13,816$:

$$\boxed{\mathbb{P}\left(12,65 \leqslant \frac{C_n}{n} \leqslant 17,02\right) \approx 0,92.}$$

L'intervalle de confiance semble meilleur qu'en première partie pour un niveau de confiance similaire. C'est normal : nous utilisions dans la première partie une inégalité de concentration peu précise (inégalité de Bienaymé-Tchebychev) alors que là nous profitions du détails de la loi. Il faut noter tout de même que l'approximation est moins bien contrôlé.

Partie III

15. (a) La variable $A_{i,m}$ compte le nombre de succès dans l'épreuve répétée m fois de Bernoulli de manière indépendante consistant en le fait d'obtenir une vignette numéro i lors d'un achat de paquet. Le succès a une probabilité de $\frac{1}{n}$ puisqu'il y a n vignettes différentes distribuées de manière équiprobable.

    Ainsi, on a bien $\boxed{A_{i,m} \hookrightarrow \mathcal{B}\left(m, \frac{1}{n}\right).}$

    (b) On a :

$$\begin{aligned}
\mathrm{Cov}(A_{1,m}, m - A_{1,m}) &= \mathrm{Cov}(A_{1,m}, m) - \mathrm{Cov}(A_{1,m}, A_{1,m}) \\
&\quad \text{(linéarité à droite)} \\
&= 0 - \mathrm{Var}(A_{1,m}) \\
&\quad \text{(car } m \text{ est une variable certaine)} \\
&= \boxed{-m \times \frac{1}{n}\left(1 - \frac{1}{n}\right).}
\end{aligned}$$

    (c) Raisonons par l'absurde et supposons que les $A_{1,m}$, ..., $A_{,nm}$ sont mutuellement indépendantes. Remarquons que l'on a : $A_{1,m} + \cdots + A_{n,m} = m$ et donc :

$$m - A_{1,m} = A_{2,m} + \cdots + A_{n,m}.$$

    Donc :

$$\mathrm{Cov}(A_{1,m}, A_{2,m} + \cdots + A_{n,m}) = \mathrm{Cov}(A_{1,m}, m - A_{1,m}) = -\frac{m}{n}\left(1 - \frac{1}{n}\right).$$

    Or par le lemme des coalitions, on a que $A_{1,m}$ et $A_{2,m} + \cdots + A_{n,m}$ sont indépendantes et donc leur covariance est nulle. D'où :

$$\boxed{0 = \frac{m}{n}\left(1 - \frac{1}{n}\right) \neq 0.}$$

    Donc, le résultat est absurde et $\boxed{\text{les } A_{1,m}, \text{ ..., } A_{,nm} \text{ ne sont pas mutuellement indépendantes.}}$

16. (a) Procédons par récurrence sur $r \in \mathbb{N}^\star$.

    - **Initialisation :** pour $r = 1$, on a :

$$\mathbb{P}\left(\bigcup_{i=1}^{1} E_i\right) = \mathbb{P}\left(E_1\right)$$

et :

$$\sum_{i=1}^{1} \mathbb{P}(E_i) = \mathbb{P}(E_1).$$

Donc les quantités sont égales et *a fortiori* :

$$\boxed{\mathbb{P}\left(\bigcup_{i=1}^{1} E_i\right) \leqslant \sum_{i=1}^{1} \mathbb{P}(E_i).}$$

- **Hérédité :** soit $r \in \mathbb{N}^\star$. On suppose que :

$$\mathbb{P}\left(\bigcup_{i=1}^{r} E_i\right) \leqslant \sum_{i=1}^{r} \mathbb{P}(E_i).$$

On a alors :

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{i=1}^{r+1} E_i\right) &\leqslant \mathbb{P}\left(\left[\bigcup_{i=1}^{r} E_i\right] \cup E_{r+1}\right) \\
&\leqslant \mathbb{P}\left(\bigcup_{i=1}^{r} E_i\right) + \mathbb{P}(E_{r+1}) \\
&\leqslant \sum_{i=1}^{r} \mathbb{P}(E_i) + \mathbb{P}(E_{r+1}) \\
&\leqslant \boxed{\sum_{i=1}^{r+1} \mathbb{P}(E_i).}
\end{aligned}$$

(b) $C_n > m$ signifie que l'album n'est pas complet au bout de m achats, donc qu'au moins une des vignettes n'a pas été obtenue. On peut écrire cela sous la forme :

$$[C_n > m] = \bigcup_{i=1}^{n} [A_{i,m} = 0].$$

Avec la question précédente, on obtient :

$$\begin{aligned}
\mathbb{P}(C_n > m) &= \mathbb{P}\left(\bigcup_{i=1}^{n} [A_{i,m} = 0]\right) \\
&\leqslant \sum_{i=1}^{n} \mathbb{P}(A_{i,m} = 0) \\
&\leqslant \sum_{i=1}^{n} \left(1 - \frac{1}{n}\right)^m \\
&\leqslant \boxed{n\left(1 - \frac{1}{n}\right)^m.}
\end{aligned}$$

On peut ensuite remarquer que pour tout $x \in \mathbb{R}$:

$$1 - x \leqslant \mathrm{e}^{-x}$$

par convexité de exp. Donc :

$$\boxed{n\left(1 - \frac{1}{n}\right)^m \leqslant n\left(\mathrm{e}^{-\frac{1}{n}}\right)^m \leqslant n\mathrm{e}^{-\frac{m}{n}}.}$$

(c) $\boxed{\text{Il y a \textit{vraisemblablement} une erreur d'énoncé.}}$
Voici *a priori*, ce qui est attendu :
Soit $c > 0$. Si $cn\ln(n)$ est entier, il suffit de poser $m = cn\ln(n)$ dans l'inégalité précédente.
Sinon, puisque C_n est à valeurs dans $\mathbb{N}^\star$ (car $n \geqslant 2$), on a :

$$\mathbb{P}(C_n > cn\ln(n)) = \mathbb{P}(C_n \geqslant \lfloor cn\ln(n)\rfloor + 1).$$

On pose alors $m = \lfloor cn\ln(n)\rfloor + 1$ dans l'inégalité précédente. On obtient :

$$\mathbb{P}(C_n > \lfloor cn\ln(n)\rfloor + 1) \leqslant n\mathrm{e}^{-\frac{\lfloor cn\ln(n)\rfloor+1}{n}} \leqslant n\mathrm{e}^{-\frac{cn\ln(n)}{n}} \leqslant n^{1-c}.$$

Et malheureusement, il faudrait $\mathbb{P}(C_n \geqslant \lfloor cn\ln(n)\rfloor + 1)$ pour conclure...

(d) On a pour tout $x \in \mathbb{R}$:

$$\mathbb{P}(V_n > x) = \mathbb{P}\left(\frac{C_n}{n} - \ln(n) > x\right) = \mathbb{P}\left(\frac{C_n}{n} > x + \ln(n)\right) = \mathbb{P}\left(C_n > \left(\frac{x}{\ln(x)} + 1\right) n\ln(n)\right).$$

On pose $c = \frac{x}{\ln(n)} + 1$. Pour $x > -\ln(n)$, on a $c > 0$ et la question précédente s'applique :

$$\boxed{\mathbb{P}(V_n > x) \leqslant n^{1-\left(\frac{x}{\ln(n)}+1\right)} = n^{-\frac{x}{\ln(x)}} = \mathrm{e}^{-\frac{x}{\ln(n)}\ln(n) = \mathrm{e}^{-x}}.}$$

17. (a) **Remarque :** encore une erreur d'énoncé ! Il faut lire « loi conditionnelle de $\tilde{A}_i$ ».
La justification est la même que pour A_i et on a bien $\tilde{A}_i \hookrightarrow \mathcal{B}\left(p, \frac{1}{n}\right)$ puisqu'on a un comptage de succès d'une épreuve de Bernoulli répétée de manière indépendante.
Ensuite, pour $k \in \mathbb{N}$, appliquons la formule des probabilités totales avec le s.c.e. $\{[N = p]\}_{p\in\mathbb{N}}$:

$$\begin{aligned}
\mathbb{P}(\tilde{A}_i = k) &= \sum_{p=0}^{+\infty} \underbrace{\mathbb{P}_{[N=p]}(\tilde{A}_i = k)}_{=0 \text{ si } k>p} \mathbb{P}(N = p) \\
&= \sum_{p=k}^{+\infty} \binom{p}{k} \frac{1}{n^k}\left(1 - \frac{1}{n}\right)^{p-k} \mathrm{e}^{-\lambda}\frac{\lambda^p}{p!} \\
&= \sum_{p=k}^{+\infty} \frac{p!}{k!(p-k)!}\frac{1}{n^k}\left(1 - \frac{1}{n}\right)^{p-k} \mathrm{e}^{-\lambda}\frac{\lambda^p}{p!} \\
&\underset{q=p-k}{=} \sum_{q=0}^{+\infty} \frac{1}{k!q!n^k}\left(1 - \frac{1}{n}\right)^{q} \mathrm{e}^{-\lambda}\lambda^{q+k} \\
&= \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!n^k}\sum_{q=0}^{+\infty} \frac{1}{q!}\left(1 - \frac{1}{n}\right)^{q}\lambda^q \\
&= \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!n^k}\mathrm{e}^{\left(1-\frac{1}{n}\right)\lambda} \\
&= \boxed{\frac{1}{k!}\left(\frac{\lambda}{n}\right)^k \mathrm{e}^{-\frac{\lambda}{n}}.}
\end{aligned}$$

Et donc $\boxed{\tilde{A}_i \hookrightarrow \mathcal{P}\left(\frac{\lambda}{n}\right).}$

(b) Sachant $N = p$ et $k_1 + \cdots + k_n = p$, l'événement $[\tilde{A}_1 = k_1] \cap \cdots \cap [\tilde{A}_n = k_n]$ correspond à :
- k_1 vignettes parmi les p ;
- k_2 vignettes parmi les $p - k_1$ restantes ;
- etc

sur un total de n^p combinaisons. On a donc :

$$\begin{aligned}
&\mathbb{P}_{[N=p]}\left([\tilde{A}_1 = k_1] \cap \cdots \cap [\tilde{A}_n = k_n]\right) \\
= &\ \frac{\binom{p}{k_1} \times \binom{p-k_1}{k_2} \times \cdots \times \binom{p-k_1-k_2-\ldots-k_{n-1}}{k_n}}{n^p} \\
= &\ \frac{1}{n^p} \times \frac{p!}{k_1!(p-k_1)!} \times \frac{(p-k_1)!}{k_2!(p-k_1-k_2)!} \times \cdots \times \frac{(p-k_1-k_2-\ldots-k_{n-1})!}{(p-k_1-k_2-\ldots-k_{n-1}-k_n)!k_n!} \\
= &\ \boxed{\frac{p!}{n^p k_1! k_2! \cdots k_n!}.}
\end{aligned}$$

(c) Soit $(k_1, \ldots, k_n) \in \mathbb{R}^n$ tel que $k_1 + \cdots + k_n = p$. On a donc :

$$[\tilde{A}_1 = k_1] \cap \cdots \cap [\tilde{A}_n = k_n] \subset [N = p]$$

Puis :

$$\begin{aligned}
\mathbb{P}([\tilde{A}_1 = k_1] \cap \cdots \cap [\tilde{A}_n = k_n]) &= \mathbb{P}([\tilde{A}_1 = k_1] \cap \cdots \cap [\tilde{A}_n = k_n] \cap [N = p]) \\
&= \mathbb{P}_{[N=p]}([\tilde{A}_1 = k_1] \cap \cdots \cap [\tilde{A}_n = k_n]) P(N = p) \\
&= \frac{p!}{n^p k_1! k_2! \cdots k_n!} \times \frac{\lambda^p}{p!} \mathrm{e}^{-\lambda} \\
&= \frac{\mathrm{e}^{-\lambda}}{k_1! \cdots k_n!} \left(\frac{\lambda}{n}\right)^{k_1 + \cdots + k_n} \\
&= \boxed{\prod_{i=1}^{n} \frac{\mathrm{e}^{-\frac{\lambda}{n}} \left(\frac{\lambda}{n}\right)^{k_i}}{k_i!}.}
\end{aligned}$$

(d) Comme $\tilde{A}_i$ suit la loi de Poisson de paramètre $\frac{\lambda}{n}$, on a :

$$\mathbb{P}(\tilde{A}_i = k) = \frac{\mathrm{e}^{-\frac{\lambda}{n}} \left(\frac{\lambda}{n}\right)^k}{k!}.$$

Et donc, d'après la question précédente, on a :

$$\boxed{\mathbb{P}([\tilde{A}_1 = k_1] \cap \cdots \cap [\tilde{A}_n = k_n]) = \prod_{i=1}^{n} \mathbb{P}(\tilde{A}_i = k_i).}$$

C'est la caractérisation de variables indépendantes pour des lois discrètes à valeurs dans $\mathbb{N}^\star$.

Donc $\boxed{\text{les } \tilde{A}_i \text{ sont mutuellement indépendantes.}}$

Malgré les apparences, ce résultat n'est pas contradictoire avec le résultat sur la non-indépendance des A_i. En effet, les A_i correspondaient à un nombre d'achats totals fixé. En conséquence, on pouvait toujours déduire A_n des A_1 à A_{n-1}.

Au contraire, le nombre d'achats ici n'est pas fixé mais suit une loi bien choisie. Apparemment, en choisissant bien cette loi, on ne peut absolument rien apprendre de A_n à partir des valeurs de A_1 à A_{n-1}.

18. On a : $D_n = \bigcap_{i=1}^{n} \left[\tilde{A}_i \geqslant 1\right]$ puisque avoir toute la collection signifie avoir au moins une vignette de chaque type. Par indépendance mutuelle, on obtient :

$$\mathbb{P}(D_n) = \prod_{i=1}^{n} \mathbb{P}(\tilde{A}_i \geqslant 1) = \prod_{i=1}^{n} \left(1 - \mathbb{P}(\tilde{A}_i = 0)\right) = \prod_{i=1}^{n} \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right) = \boxed{\left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right)^n.}$$

19. (a) On utilise la formule des probabilités totales avec le s.c.e. $\{[N = p]\}_{p \in \mathbb{N}}$. On a :

$$\mathbb{P}(D_n) = \sum_{p=0}^{+\infty} \mathbb{P}_{N=p}(D_n) \mathbb{P}(N = p) = \boxed{\sum_{p=0}^{+\infty} \mathbb{P}(C_n \leqslant p) \mathbb{P}(N = p)}$$

en utiilisant la formule donnée par l'énoncé.

(b) On a :

$$\sum_{p=k_1}^{k_2} \mathbb{P}(N=p) = \mathbb{P}(\lfloor \lambda - a \rfloor \leqslant N \leqslant \lfloor \lambda + a \rfloor + 1).$$

Pour tout $x \in \mathbb{R}$, on a $\lfloor x \rfloor \leqslant x$ donc : $[\lfloor \lambda - a \rfloor \leqslant N] \supset [\lambda - a \leqslant N]$.
Et pour tout $x \in \mathbb{R}$, on a $x < \lfloor x \rfloor + 1$ donc : $[N \leqslant \lambda + a] \subset [N < \lfloor \lambda + a \rfloor + 1] \subset [N \leqslant \lfloor \lambda + a \rfloor + 1]$.
On obtient donc :

$$\sum_{p=k_1}^{k_2} \mathbb{P}(N=p) \geqslant \mathbb{P}(\lambda - a \leqslant N \leqslant \lambda + a) = 1 - \mathbb{P}(|N-\lambda| > a).$$

N suit une loi de Poisson donc admet une variance et une espérance. On a d'après l'inégalité de Bienaymé-Tchebychev :

$$\forall \epsilon > 0,\ \mathbb{P}(|N - \underbrace{E(N)}_{=\lambda}| \geqslant \epsilon) \leqslant \frac{\overbrace{\mathrm{Var}(V)}^{=\lambda}}{\epsilon^2}.$$

En particulier pour $\epsilon = a$:

$$\mathbb{P}(|N-\lambda| > a) \leqslant \mathbb{P}(|N-\lambda| \geqslant a) \leqslant \frac{\lambda}{a^2}.$$

Et donc :

$$\boxed{\sum_{p=k_1}^{k_2} \mathbb{P}(N=p) \geqslant 1 - \frac{\lambda}{a^2}.}$$

(c) On a clairement :

$$\mathbb{P}(D_n) = \sum_{p=0}^{+\infty} \mathbb{P}(C_n \leqslant p)\mathbb{P}(N=p) \geqslant \sum_{p=k_1}^{k_2} \mathbb{P}(C_n \leqslant p)\mathbb{P}(N=p).$$

Or pour $p \geqslant k_1$, on a :

$$[C_n \leqslant k_1] \subset [C_n \leqslant p].$$

Donc :

$$\mathbb{P}(D_n) \geqslant \sum_{p=k_1}^{k_2} \mathbb{P}(C_n \leqslant k_1)\mathbb{P}(N=p) \geqslant \mathbb{P}(C_n \leqslant k_1) \sum_{p=k_1}^{k_2} \mathbb{P}(N=p) \geqslant \boxed{\mathbb{P}(C_n \leqslant k_1)\left(1 - \frac{\lambda}{a^2}\right).}$$

Pour la seconde inégalité, on part de :

$$\mathbb{P}(D_n) = \sum_{p=0}^{+\infty} \mathbb{P}(C_n \leqslant p)\mathbb{P}(N=p) = \sum_{p=0}^{k_2} \mathbb{P}(C_n \leqslant p)\mathbb{P}(N=p) + \sum_{p=k_2+1}^{+\infty} \mathbb{P}(C_n \leqslant p)\mathbb{P}(N=p).$$

Comme précédemment, on a :

$$\sum_{p=0}^{k_2} \mathbb{P}(C_n \leqslant p)\mathbb{P}(N=p) \leqslant \mathbb{P}(C_n \leqslant k_2) \underbrace{\sum_{p=0}^{k_2} \mathbb{P}(N=p)}_{\leqslant 1} \leqslant \mathbb{P}(C_n \leqslant k_2).$$

Et on a également puisque $\mathbb{P}(C_n \leqslant p) \leqslant 1$:

$$\sum_{p=k_2+1}^{+\infty} \mathbb{P}(C_n \leqslant p)\mathbb{P}(N=p) \leqslant \sum_{p=k_2+1}^{+\infty} \mathbb{P}(N=p).$$

Or :

$$\sum_{p=k_2+1}^{+\infty} \mathbb{P}(N=p) = 1 - \sum_{0}^{k_2} \mathbb{P}(N=p) \leqslant 1 - \sum_{p=k_1}^{k_2} \mathbb{P}(N=p) \leqslant \frac{\lambda}{a^2}.$$

En remettant bout à bout, on obtient bien :

$$\boxed{\mathbb{P}(D_n) \leqslant \mathbb{P}(C_n \leqslant k_2) + \frac{\lambda}{a^2}.}$$

20. (a) Posons comme le suggère l'énoncé : $\lambda = n\ln(n) + c_n n$ et $a = n^{2/3}$.
D'après la question précédente, on a :

$$\mathbb{P}(C_n \leqslant k_1)\left(1 - \frac{\lambda}{a^2}\right) \leqslant \mathbb{P}(D_n) \leqslant \mathbb{P}(C_n \leqslant k_2) + \frac{\lambda}{a^2}.$$

Or on a :

$$\begin{aligned}\mathbb{P}(C_n \leqslant k_1) &= \mathbb{P}(C_n \leqslant \lfloor \lambda - a \rfloor) \geqslant \mathbb{P}(C_n \leqslant \lambda - a - 1) \\ &\geqslant \mathbb{P}(nV_n + n\ln(n) \leqslant n\ln(n) + c_n n - n^{2/3} - 1) \geqslant \mathbb{P}\left(V_n \leqslant c_n - \frac{1}{n^{1/3}} - \frac{1}{n}\right).\end{aligned}$$

On a également :

$$1 - \frac{\lambda}{a^2} = 1 - \frac{n\ln(n) + c_n n}{(n^{2/3})^2} = 1 - \frac{\ln(n) + c_n}{n^{1/3}}.$$

Donc ;

$$\boxed{\mathbb{P}\left(V_n \leqslant c_n - \frac{1}{n^{1/3}} - \frac{1}{n}\right)\left(1 - \frac{\ln(n) + c_n}{n^{1/3}}\right) \leqslant \mathbb{P}(D_n).}$$

De manière tout à fait similaire, on trouve :

$$\boxed{\mathbb{P}(D_n) \leqslant \mathbb{P}\left(V_n \leqslant c_n + \frac{1}{n^{1/3}} + \frac{1}{n}\right) + \frac{\ln(n) + c_n}{n^{1/3}}.}$$

Puis avec une question précédente, on a :

$$\mathbb{P}(D_n) = \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right)^n = \left(1 - \mathrm{e}^{-\frac{n\ln(n) + c_n n}{n}}\right)^n = \boxed{\left(1 - \frac{\mathrm{e}^{-c_n}}{n}\right)^n.}$$

Et on a donc bien l'inégalité voulue.

(b) Pour $x \in \mathbb{R}$, on pose $c_n = x + \frac{1}{n^{1/3}} + \frac{1}{n}$. On a $c_n \xrightarrow[n\to+\infty]{} x$.
Comme $\frac{\ln(n)+c_n}{n^{1/3}} \xrightarrow[n\to+\infty]{} 0$ par croissance comparée, on a pour n suffisamment grand :

$$1 - \frac{\ln(n) + c_n}{n^{1/3}} > 0$$

On a donc :

$$\mathbb{P}\left(V_n \leqslant \underbrace{c_n - \frac{1}{n^{1/3}} - \frac{1}{n}}_{=x}\right) \leqslant \underbrace{\frac{\left(1 - \frac{\mathrm{e}^{-c_n}}{n}\right)^n}{1 - \frac{\ln(n)+c_n}{n^{1/3}}}}_{\to \mathrm{e}^{-\mathrm{e}^{-x}}}.$$

De même, en posant $c_n = x - \frac{1}{n^{1/3}} - \frac{1}{n}$, on trouve :

$$\mathbb{P}\left(V_n \leqslant \underbrace{c_n + \frac{1}{n^{1/3}} + \frac{1}{n}}_{=x}\right) \geqslant \underbrace{\left(1 - \frac{\mathrm{e}^{-c_n}}{n}\right)^n - \frac{\ln(n) + c_n}{n^{1/3}}}_{\to \mathrm{e}^{-\mathrm{e}^{-x}}}.$$

Par encadrement, on trouve :

$$\boxed{F_{V_n}(x) \xrightarrow[n\to+\infty]{} \mathrm{e}^{-\mathrm{e}^{-x}}.}$$

Donc $\boxed{(V_n) \text{ converge en loi vers la loi de Gumbel.}}$