# CC2 2023/2024 – Durée 1h

*Les documents et appareils électroniques (calculatrice, téléphone, ordinateur, ...) sont interdits. Il est sous-entendu que les estimateurs, intervalles de confiance et tests demandés doivent être non triviaux.* ***Toutes les réponses doivent être justifiées****.*

## Exercice 1 - Nombre de pannes annuelles dans une concession

Une gérante de concession automobile cherche à planifier son budget réparations. Pour ce faire, elle collecte, sur les $n$ dernières années, le nombre de pannes survenues (par an) dans son parc. Le nombre de pannes annuel dans un parc automobile est bien modélisé par une loi de Poisson de paramètre $\lambda > 0$, et on suppose que les nombres de pannes par années sont indépendants.

1. Proposer un modèle pour cette expérience.
2. Donner la densité associée à ce modèle.
3. Proposer un estimateur de $\lambda$, basé sur la méthode des moments (on le notera $\hat{\lambda}_n$).
4. Montrer que $\hat{\lambda}_n$ est consistant.
5. Montrer que $\hat{\lambda}_n$ est asymptotiquement normal.
6. En déduire que $\sqrt{n}(\sqrt{\hat{\lambda}_n} - \sqrt{\lambda}) \rightsquigarrow \mathcal{N}(0, 1/4)$.
7. En déduire un intervalle de confiance **asymptotique** sur $\sqrt{\lambda}$, de niveau de confiance $95\%$, et du type $[0; \hat{s}_+]$.
8. En déduire un intervalle de confiance **asymptotique** sur $\lambda$, de niveau de confiance $95\%$, et du type $[0; \hat{\lambda}_+]$.
9. Montrer que $E_\lambda((\hat{\lambda}_n - \lambda)^2) = \lambda/n$.
10. (**) En déduire un intervalle de confiance (**non-asymptotique**), de niveau de confiance $95\%$ sur $\lambda$, de type $[0; \hat{\lambda}_+]$.

## Solution 1 -

1. On note $X_i$ le nombre de pannes survenus dans le parc auto au cours de la $i$-ème dernière année. On peut supposer que les $X_i$ sont i.i.d. de loi $\mathcal{P}(\lambda)$, pour un paramètre $\lambda > 0$ inconnu. Le modèle s'écrit alors

$$\left( (\mathbb{N}^n, \left(\mathcal{P}(\lambda)^{\otimes n}\right)_{\lambda > 0} \right).$$

2. La densité est donnée par, pour $x_1, \dots, x_n \in \mathbb{N}$,

$$f_\lambda(x_{1:n}) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i} \prod_{i=1}^{n} \frac{1}{x_i!}.$$

3. Comme $E_\lambda(X_1) = \lambda$, l'estimateur par moments $\hat{\lambda}_n$ vérifie $\bar{X}_n = \hat{\lambda}_n$.
4. Comme $E_\lambda(|X_1|) = \lambda < +\infty$, la loi des grands nombres donne

$$\bar{X}_n = \hat{\lambda}_n \xrightarrow[n\to+\infty]{\mathbb{P}} \lambda.$$

   $\hat{\lambda}_n$ est donc consistant.
5. Comme $\mathrm{Var}_\lambda(X_1) = \lambda < +\infty$, le théorème central limite donne

$$\sqrt{n}(\hat{\lambda}_n - \lambda) \rightsquigarrow \mathcal{N}(0, \mathrm{Var}_\lambda(X_1)),$$

   avec $\mathrm{Var}_\lambda(X_1) = \lambda$. $\hat{\lambda}_n$ est donc asymptotiquement normal.
6. Comme $g : x \mapsto \sqrt{x}$ est dérivable en $\lambda$, la méthode Delta donne

$$\sqrt{n}(\sqrt{\hat{\lambda}_n} - \sqrt{\lambda}) \rightsquigarrow g'(\lambda)\mathcal{N}(0, \lambda) \sim \frac{1}{2\sqrt{\lambda}}\mathcal{N}(0, \lambda) \sim \mathcal{N}(0, 1/4).$$

7. D'après la question précédente on déduit

$$2\sqrt{n}(\sqrt{\hat{\lambda}_n} - \sqrt{\lambda}) \rightsquigarrow \mathcal{N}(0, 1).$$

   Soit $q$ le quantile d'ordre $95\%$ d'une loi $\mathcal{N}(0, 1)$. On a alors

$$P_\lambda\left(\left\{2\sqrt{n}(\sqrt{\lambda} - \sqrt{\hat{\lambda}_n}) \leq q\right\}\right) \xrightarrow[n\to+\infty]{} \mathbb{P}\left(\mathcal{N}(0,1) \leq q\right) = 95\%.$$

   Or, pour $\lambda \geq 0$,

$$2\sqrt{n}(\sqrt{\lambda} - \sqrt{\hat{\lambda}_n}) \leq q \quad \Leftrightarrow \quad 0 \leq \sqrt{\lambda} \leq \sqrt{\hat{\lambda}_n} + \frac{q}{2\sqrt{n}}.$$

   On en déduit que

$$\left[0; \sqrt{\hat{\lambda}_n} + \frac{q}{2\sqrt{n}}\right]$$

   est un intervalle de niveau de confiance asymptotique $95\%$.

8. Comme $x \mapsto x^2$ est croissante sur $\mathbb{R}^+$ (et $\lambda > 0$), on déduit de la question précédente que
$$\left[0; \left(\sqrt{\hat{\lambda}_n} + \frac{q}{2\sqrt{n}}\right)^2\right]$$
est un intervalle de niveau de confiance asymptotique $95\%$ sur $\lambda$.
9. On a $E_\lambda(\hat{\lambda}_n) = \lambda$, et $\mathrm{Var}_\lambda(\hat{\lambda}_n) = \mathrm{Var}_\lambda(X_1)/n = \lambda/n$. Une décomposition biais-variance donne alors
$$E_\lambda((\hat{\lambda}_n - \lambda)^2) = \mathrm{Var}_\lambda(\hat{\lambda}_n) + (E_\lambda(\hat{\lambda}_n) - \lambda)^2 = \lambda/n.$$
10. On utilise l'inégalité de Markov. Pour $t > 0$,
$$P_\lambda\left(\left\{\lambda - \hat{\lambda}_n > t\right\}\right) \leq P_\lambda\left(\left\{(\hat{\lambda}_n - \lambda)^2 > t^2\right\}\right) \leq \frac{E_\lambda\left((\hat{\lambda}_n - \lambda)^2\right)}{t^2} = \frac{\lambda}{nt^2}.$$
Posons $t_{5\%} = \sqrt{\frac{\lambda}{n5\%}}$, de sorte que
$$P_\lambda\left(\left\{\lambda > \hat{\lambda}_n + t_{5\%}\right\}\right) \geq 1 - 5\% = 95\%,$$
d'après ce qui précède. On peut réécrire, pour $\lambda \geq 0$,
$$\lambda \leq \hat{\lambda}_n + t_{5\%} \quad \Leftrightarrow \quad \lambda \leq \hat{\lambda}_n + \sqrt{\frac{\lambda}{n5\%}}.$$
Cette inéquation peut se réécrire
$$\left(\sqrt{\lambda} - \frac{1}{2\sqrt{n5\%}}\right)^2 \leq \hat{\lambda}_n + \frac{1}{n20\%},$$
ce qui mène à
$$\sqrt{\lambda} \leq \frac{1}{2\sqrt{n5\%}} + \sqrt{\hat{\lambda}_n + \frac{1}{n20\%}},$$
et donc à
$$\lambda \leq \left(\frac{1}{2\sqrt{n5\%}} + \sqrt{\hat{\lambda}_n + \frac{1}{n20\%}}\right)^2.$$
Un intervalle au niveau de confiance $95\%$ pour $\lambda$ est donc
$$\left[0; \left(\frac{1}{2\sqrt{n5\%}} + \sqrt{\hat{\lambda}_n + \frac{1}{n20\%}}\right)^2\right].$$